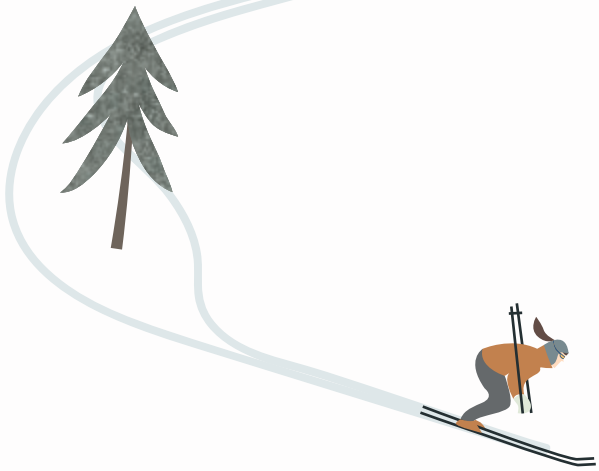
Organ preservation in rectal cancer

ORGAN PRESERVATION IN RECTAL CANCER - CAPITA SELECTA



HESTER ELINE HAAK

ORGAN PRESERVATION IN RECTAL CANCER
Capita Selecta

Hester Eline Haak

The work described in this thesis was performed at the Netherlands Cancer Institute – Antoni van leeuwenhoek, Amsterdam, the Netherlands

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing: Ridderprint
Cover and layout: Elisa Calamita, persoonlijkproefschrift.nl
ISBN: 978-94-6416-834-1

The publication of this thesis was financially supported by the Netherlands Cancer Institute, Maastricht University, ChipSoft and Erbe Nederland B.V.

ORGAN PRESERVATION IN RECTAL CANCER

Capita Selecta

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,

op gezag van de Rector Magnificus, Prof. dr. Pamela Habibović

volgens het besluit van het College van Decanen,

in het openbaar te verdedigen

op dinsdag 7 juni 2022 om 10.00 uur

door

Hester Eline Haak

geboren op 22 februari 1991 te ‘s-Gravenhage

Promotores
Prof. dr. G.L. Beets
Prof. dr. R.G.H. Beets-Tan

Copromotor

Dr. M. Maas	Antoni van Leeuwenhoek

Beoordelingscommissie

Prof. dr. H.J.T. Rutten (voorzitter)	
Prof. dr. A.P. de Bruïne	
Dr. K. Horsthuis	Amsterdam UMC, locatie VUmc
Prof. dr. C.A.M. Marijnen	Leids Universitair Medisch Centrum

TABLE OF CONTENTS

CHAPTER 1

General introduction and thesis outline

This section is based on the following book chapter:
Evaluation of tumour response after radiotherapy in rectal cancer.
Hester E. Haak, Geerard L. Beets

Baatrup G. (eds) Multidisciplinary Treatment of Colorectal Cancer. Springer, Cham. 2021:249-54

Nowadays, colorectal cancer is the 3rd most common cancer worldwide, and each year ±1.8 million new patients are diagnosed in both men and women.[1] Risk factors are old age, male gender, obesity, physical inactivity, poor diets, alcohol and smoking.[2] In The Netherlands, the incidence of colorectal cancer increased due to the implementation of the bowel screening program in 2014, and now lies around approximately 12.000 new patients per year of which one third is located in the rectum.[3] The treatment and outcomes of colon cancer are different from rectal cancer and thus need to be regarded as different entities, that need to be evaluated separately. Recently, a modern definition of the sigmoid and rectum was adopted, based on the so-called "sigmoid take-off", with the aim to provide a more reproducible definition in order to standardize treatment and inclusion in national and international registrations and trials.[4, 5]

Primary staging and treatment

In the past, approximately 30% of patients with rectal cancer experienced a local recurrence, usually due to suboptimal surgical technique with blunt resection of the tumour. These local recurrences often required extensive surgery or were irresectable.[6, 7] Due to several advances in treatment the local recurrence rate is now down to 3-5%.[8, 9] First, the introduction of the total mesorectal excision (TME) reduced the risk of local recurrence significantly.[7] Standard TME contains removal of the mesorectum with a sharp dissection between the visceral and parietal layers including the whole mesorectal envelope with regional vessels and lymph nodes, which minimizes the risk for a recurrence in tumour deposits or remaining lymph nodes. Second, the Dutch TME trial showed that pre-operative radiotherapy reduces the risk of local recurrence from 26-27% in the non-irradiated group to 9-11% in the irradiated group after TME.[8] In addition, Sauer et al. showed that the risk of local recurrence further reduced from 13 to 6% in patients with locally advanced tumours who were randomized to preoperative chemoradiation (CRT) compared to postoperative CRT.[10] Third, substantial improvements in pre-treatment imaging by the introduction of high quality magnetic resolution imaging (MRI) were made which led to improved risk stratification of rectal tumours.[11]

Primary staging is mostly done with digital rectal examination, endoscopy, and MRI. After primary staging, the decision for treatment is discussed in a multidisciplinary meeting.[12] In Europe, rectal cancer treatment is based on three risk groups: low risk, intermediate risk and high risk.[13] In short, patients with a low risk tumour (cT1-3abN0) require direct surgery with standard TME. Patients with an intermediate risk tumour (very distal cT2-3ab without mesorectal fascia invasion(MRF-)N0 or cT1-3(MRF-)N1) require a short course of radiotherapy with immediate TME within one week after radiation. High risk patients (cT3cd, cT4, cN2 or suspect extramesorectal nodes, threatened or invaded MRF, extramural vascular invasion) need to undergo a long course of neoadjuvant CRT and a long waiting interval (6-12 weeks) followed by TME.

Despite the several improvements, TME and CRT have its downsides. TME is related with a high morbidity, impaired anorectal and urogenital dysfunction, receiving a permanent colostomy in distal tumours and outcomes are worse in older or frail patients.[7, 14] In addition, CRT may cause severe acute toxicity grade 3 in 21% of the patients with diarrhea in 10% and radiation dermatitis in 12% as the most frequent complications.[15, 16] Moreover, although the incidence and local recurrence decreased, mortality did not improve.[17] Overall, the 5-year survival is 67% but varies according to different risk groups with 93% in low risk tumours and 17% in case of metastatic disease.[18]

Response assessment after neoadjuvant treatment

In the past, restaging of rectal cancer was mainly performed to rule out progression and to assess whether the original surgical plan was still valid. Occasionally, a very good response in locally advanced tumours allowed for a less extensive resection when a cT4 tumour showed regression from the invaded organ.[19] In about 20% of the patients the tumour is completely gone, a so-called pathological complete response, which is associated with favourable long-term outcomes compared to patients who do not have a pathological complete response.[9] Over the past decade, the interest has risen to perform organ preserving therapies in patients with a clinical complete response to avoid major operation with less extensive surgery (i.e. local excision) or omission of surgery, a so-called watch-and-wait (W&W) approach. In 2004, Habr Gama et al. was the first who reported excellent outcomes of W&W patients with a 10-year overall survival of 98% and a 10-year disease free survival of 84%.[20] Since then, more research groups confirmed these excellent outcomes [21-23], and W&W patients are now being registered in the International Watch-and-Wait Database (IWWD) by participating centres worldwide. [24, 25]

The current interest in organ preservation has renewed the interest to assess the response after (C)RT, with the goal to identify patients in whom a local excision or W&W policy can be proposed as an alternative to TME. The traditional interval to surgery after the CRT used to be 6–8 weeks, and was rather arbitrarily chosen as a good compromise between allowing for a maximal downsizing effect and minimizing the risk of progression in non-responders. In order to identify possible complete responders a longer waiting interval of 8-12 weeks is advised.[26] In some patients a very good but not typical clinical complete response is seen and it is advised to perform a second evaluation 6–12 weeks later, rather than proceed to TME surgery after the first evaluation. This approach in so-called 'near complete responders' has been shown to develop into a clinical complete response at the second evaluation 6–12 weeks later in a majority of the patients.[27] Clinical assessment with digital rectal examination and endoscopy is the single most accurate modality for identification of clinical complete responders. The most commonly used endoscopy technique is a standard high-resolution endoscopy with white light. T2-weighted high-resolution MR imaging is the standard to provide morphological imaging of the luminal and extra-luminal response. The exact pattern of response and fibrosis can be helpful to identify clinical complete responders[28], although the differentiation of vital tumour within the fibrosis remains a challenge with T2-

weighted MRI.[29] The addition of diffusion-weighted (DWI) MRI has resulted in an improved performance to differentiate between patients with a clinical complete responders and those with residual tumour.[29-31] The combination of digital rectal examination, endoscopy and MRI has the highest accuracy to detect clinical complete responders with a low risk of missing residual disease.[32] The European Society of Gastrointestinal and Abdominal Radiology guidelines recommend to perform this three-modality approach when considering organ preservation (W&W) after CRT.[33]

W&W

W&W has been recently implemented in Dutch guidelines as an alternative treatment in patients who have a clinical complete responders after neoadjuvant therapy.[13] However, it is emphasized that W&W needs to be performed by expert centres with dedicated clinicians. The multicentre prospective W&W implementation study (NCT03426397), open for inclusion since 2017, provides this platform where 13 national expert centres from The Netherlands provide the W&W approach in a quality-controlled setting. Simultaneously with the cautious approach of registering all W&W patients in this platform, the interest in organ preservation has increased and organ preservation is increasingly pursued in patients with low to intermediate risk tumours who do not necessarily require neoadjuvant treatment. One option for organ preservation that already is adopted in current guidelines is to perform local excision in patients with T1 tumours. When the tumour has no adverse risk factors on histology (i.e. clear margins, well/moderately differentiated, only superficial invasion of the submucosa (sm1-2)) TME can be omitted with a low risk of lymph node metastasis. However, in case of the presence of adverse risk factors additional TME is still recommended. In these patients two alternative options can be considered: careful follow-up with salvage surgery when the residual disease becomes evident, or adjuvant CRT. Patients with low-intermediate risk tumours may also alternatively receive upfront neoadjuvant treatment to increase the possibility of organ preservation, however, with the expense of a higher risk for morbidity.[34]

The main goal of W&W is to avoid TME and a permanent stoma and to improve quality of life without compromising local recurrence and overall survival. Results of the first 1000 inclusions of the IWWD showed that oncological outcomes were good with a 5-year overall survival of 85%, a 5-year disease specific survival of 94% and only 8% risk to develop distant metastasis.[24] In total, 25% of patients developed a local regrowth which mostly occurs within two years and are located in the lumen. Frequent follow-up is necessary to identify local regrowths early in order to achieve similar long-term outcomes compared to patients who undergo a standard rectal resection. Follow-up schedules differ between centres, but most patients are intensively followed with 3-montly endoscopy and MRI during the first two years and 6-monthly thereafter. A local regrowth should not be misinterpreted as a local recurrence which implies surgical failure and has a poor prognosis[35] whereas the estimated risk of locally unsalvageable disease of local regrowths is expected to be around 1%.[24] However, some research groups report that patients with a local regrowth have a higher risk to develop distant metastasis.[24, 36, 37] There are two hypotheses for the increased risk of

metastasis in local regrowths. One is that the higher risk is due to omission of surgery, the other is that the higher risk is caused by unfavourable tumour biology. This question has not been answered yet. Another important goal of W&W is to improve functional outcomes and quality of life. Only a few studies described these outcomes. Hupkens et al. showed that after a successful W&W, the quality of life was better on several domains compared to patients treated with CRT and surgery.[38] However, CRT on its own was not without side-effects as 30% of W&W patients still had major low anterior resection syndrome symptoms compared to 67% in patients treated with CRT and surgery. Furthermore, local excision after CRT is associated with poor functional outcome, demonstrated by the study of Habr Gama et al. where patients with a near complete response who were treated with local excision and W&W had worse anorectal function and quality of life compared to patients with a clinical complete responders treated with W&W only.[39] These outcomes show that preservation of function is only achieved with a non-operative approach in near complete responders. In addition, due to the implementation of the national bowel screening program and aging of the population more older patients are diagnosed with rectal cancer, and there is little information on the oncological and functional outcomes of a W&W approach in older patients.

Aims of this thesis

The favorable oncological and functional outcomes of complete responders results in increasing interest in organ preserving therapies. These less invasive treatments have its advantages, but the approach itself introduces its own risks, complexities and management difficulties. Risks should be well balanced for each patient, whilst taking the preferences of the patient into account. Although a lot of research has been performed, several challenges remain. The following objectives are addressed in this thesis:

- To evaluate if MRI alone can accurately identify patients who have substantial residual disease after neoadjuvant chemoradiation that requires immediate surgery, and if these findings are reproducible amongst radiologists with variable levels of expertise.
- To evaluate the pooled prevalence of lymph nodes after chemoradiation according to increasing depth of residual tumour in the rectal wall and to assess the impact of post-chemoradiation lymph nodes metastases on long-term oncological outcomes.
- To evaluate the current watch-and-wait follow-up schedule and to propose improvements to make the follow-up schedule more efficient.
- To evaluate the oncological and functional outcomes of a watch-and-wait approach in older patients.
- To evaluate if distant metastasis occur later in watch-and-wait patients than in patients treated with chemoradiation and total mesorectal excision by comparing metastasis and detection.
- To give an overview of current and new imaging technologies for prediction and assessment of response.

Outline of this thesis

This thesis is divided in three parts. *Part I* evaluates current patient selection and follow-up. In *chapter 2,* we assess if radiologists with variable levels of expertise are able to accurately identify poor responders on MRI using a simplified thee-categorized response evaluation system. In *chapter 3*, we evaluate the prevalence of lymph node metastasis in patients with locally advanced rectal cancer treated with chemoradiation and total mesorectal excision according ypT-stage and evaluate oncological outcomes of ypN status in terms of disease recurrence and survival. In *chapter 4***,** we evaluate the occurrence and detection of local regrowths in a retrospective W&W cohort and propose a more efficient follow-up schedule.

Part II evaluates oncological outcomes of a watch-and-wait approach. In *chapter 5***,** we assess the outcome of a watch-and-wait approach in patients at age 75 or older. In *chapter 6*, we compare the time pattern of distant metastases in patients who followed a watch-and-wait approach after neoadjuvant chemoradiation who are registered in the International Watch-and-Wait Database to a pooled dataset of patients treated with neoadjuvant chemoradiation and total mesorectal excision.

Part III evaluates recent advanced imaging technologies. In *chapter 7,* we provide advances in MR imaging for prediction and assessment of response and give further directions for future research. In *chapter 8*, we perform a retrospective study to evaluate if the use of deep learning on endoscopic images is able to identify complete responders. Finally, *chapter 9* includes a general discussion and future perspectives.

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. Apr 15 2019;144(8):1941-1953.
2. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nature Reviews Gastroenterology & Hepatology*. 2019/12/01 2019;16(12):713-732.
3. IKNL. Incidentie darmkanker. Accessed 2021/03/07, 2021. https://iknl.nl/kankersoorten/darmkanker/registratie/incidentie
4. D'Souza N, de Neree Tot Babberich MPM, d'Hoore A, et al. Definition of the Rectum: An International, Expert-based Delphi Consensus. *Ann Surg*. Dec 2019;270(6):955-959.
5. D'Souza N, Lord A, Shaw A, et al. The sigmoid take-off: An anatomical imaging definition of the rectum validated on specimen analysis. *Eur J Surg Oncol*. Sep 2020;46(9):1668-1672.
6. Hermanek P, Wiebelt H, Staimmer D, Riedl S. Prognostic factors of rectum carcinoma--experience of the German Multicentre Study SGCRC. German Study Group Colo-Rectal Carcinoma. *Tumori*. May-Jun 1995;81(3 Suppl):60-4.
7. Heald RJ, Ryall RD. Recurrence and survival after total mesorectal excision for rectal cancer. *Lancet*. Jun 28 1986;1(8496):1479-82.
8. Kapiteijn E, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *N Engl J Med*. Aug 30 2001;345(9):638-46.
9. Maas M, Nelemans PJ, Valentini V, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol*. Sep 2010;11(9):835-44.
10. Sauer R, Becker H, Hohenberger W, et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. *N Engl J Med*. Oct 21 2004;351(17):1731-40.
11. Taylor FG, Quirke P, Heald RJ, et al. Preoperative high-resolution magnetic resonance imaging can identify good prognosis stage I, II, and III rectal cancer best managed by surgery alone: a prospective, multicenter, European study. *Ann Surg*. Apr 2011;253(4):711-9.
12. Engelen SM, Maas M, Lahaye MJ, et al. Modern multidisciplinary treatment of rectal cancer based on staging with magnetic resonance imaging leads to excellent local control, but distant control remains a challenge. *Eur J Cancer*. Jul 2013;49(10):2311-20.
13. Specialisten RFM. Primaire behandeling rectumcarcinoom. Accessed 2021/03/07, 2021. https://richtlijnendatabase.nl/richtlijn/colorectaal_carcinoom_crc/primaire_behandeling_rectumcarcinoom_bij_crc.html
14. Rutten HJ, Den Dulk M, Lemmens VE, Van De Velde CJ, Marijnen CA. Controversies of total mesorectal excision for rectal cancer in elderly patients. *The Lancet Oncology*. 2008;9(5):494-501.
15. Swellengrebel HA, Marijnen CA, Verwaal VJ, et al. Toxicity and complications of preoperative chemoradiotherapy for locally advanced rectal cancer. *Br J Surg*. Mar 2011;98(3):418-26.
16. Pucciarelli S, Del Bianco P, Efficace F, et al. Patient-reported outcomes after neoadjuvant chemoradiotherapy for rectal cancer: a multicenter prospective observational study. *Ann Surg*. Jan 2011;253(1):71-7.

17. van Gijn W, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. *Lancet Oncol*. Jun 2011;12(6):575-82. doi:10.1016/s1470-2045(11)70097-3
18. IKNL. Endeldarmkanker: Overleving, Stadium bij diagnose, TNM 7e editie (2010-2016). Accessed 2021/05/07, 2021. https://iknl.nl/nkr-cijfers
19. Vliegen RF, Beets GL, Lammering G, et al. Mesorectal fascia invasion after neoadjuvant chemotherapy and radiation therapy for locally advanced rectal cancer: accuracy of MR imaging for prediction. *Radiology*. Feb 2008;246(2):454-62.
20. Habr-Gama A, Perez RO, Nadalin W, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg*. Oct 2004;240(4):711-7; discussion 717-8.
21. Appelt AL, Pløen J, Harling H, et al. High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. *The Lancet Oncology*. 2015/08/01/ 2015;16(8):919-927.
22. Renehan AG, Malcomson L, Emsley R, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *The Lancet Oncology*. 2016;17(2):174-183.
23. Martens MH, Maas M, Heijnen LA, et al. Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. *J Natl Cancer Inst*. Dec 2016;108(12)
24. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
25. Fernandez LM, São Julião GP, Figueiredo NL, et al. Conditional recurrence-free survival of clinical complete responders managed by watch and wait after neoadjuvant chemoradiotherapy for rectal cancer in the International Watch & Wait Database: a retrospective, international, multicentre registry study. *Lancet Oncol*. Jan 2021;22(1):43-50. doi:10.1016/s1470-2045(20)30557-x
26. Sloothaak DA, Geijsen DE, van Leersum NJ, et al. Optimal time interval between neoadjuvant chemoradiotherapy and surgery for rectal cancer. *Br J Surg*. Jun 2013;100(7):933-9.
27. Hupkens BJP, Maas M, Martens MH, et al. Organ Preservation in Rectal Cancer After Chemoradiation: Should We Extend the Observation Period in Patients with a Clinical Near-Complete Response? *Ann Surg Oncol*. Jan 2018;25(1):197-203.
28. Lambregts DMJ, Delli Pizzi A, Lahaye MJ, et al. A Pattern-Based Approach Combining Tumor Morphology on MRI With Distinct Signal Patterns on Diffusion-Weighted Imaging to Assess Response of Rectal Tumors After Chemoradiotherapy. *Dis Colon Rectum*. Mar 2018;61(3):328-337.
29. Lambregts DMJ, Boellaard TN, Beets-Tan RGH. Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging: a pictorial review. *Insights into imaging*. 2019;10(1)
30. van der Paardt MP, Zagers MB, Beets-Tan RG, Stoker J, Bipat S. Patients who undergo preoperative chemoradiotherapy for locally advanced rectal cancer restaged by using diagnostic MR imaging: a systematic review and meta-analysis. *Radiology*. Oct 2013;269(1):101-12.
31. Lambregts DMJ, Van Heeswijk MM, Delli Pizzi A, et al. Diffusion-weighted MRI to assess response to chemoradiotherapy in rectal cancer: main interpretation pitfalls and their use for teaching. *Eur Radiol*. 2017 Oct;27(10):4445-4454.

32. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol*. Nov 2015;22(12):3873-80.
33. Beets-Tan RGH, Lambregts DMJ, Maas M, et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *Eur Radiol*. Apr 2018;28(4):1465-1475.
34. Perez RO, Habr-Gama A, Sao Juliao GP, Proscurshim I, Scanavini Neto A, Gama-Rodrigues J. Transanal endoscopic microsurgery for residual rectal cancer after neoadjuvant chemoradiation therapy is associated with significant immediate pain and hospital readmission rates. *Dis Colon Rectum*. May 2011;54(5):545-51.
35. Heald RJ, Beets G, Carvalho C. Report from a consensus meeting: response to chemoradiotherapy in rectal cancer - predictor of cure and a crucial new choice for the patient: on behalf of the Champalimaud 2014 Faculty for 'Rectal cancer: when NOT to operate'. *Colorectal Dis*. May 2014;16(5):334-7.
36. Socha J, Kępka L, Michalski W, Paciorek K, Bujko K. The risk of distant metastases in rectal cancer managed by a watch-and-wait strategy – A systematic review and meta-analysis. *Radiother Oncol*. 2020;144:1-6.
37. Smith JJ, Strombom P, Chow OS, et al. Assessment of a Watch-and-Wait Strategy for Rectal Cancer in Patients With a Complete Response After Neoadjuvant Therapy. *JAMA oncology*. Jan 10 2019:e185896.
38. Hupkens BJP, Martens MH, Stoot JH, et al. Quality of Life in Rectal Cancer Patients After Chemoradiation: Watch-and-Wait Policy Versus Standard Resection - A Matched-Controlled Study. *Dis Colon Rectum*. Oct 2017;60(10):1032-1040.
39. Habr-Gama A, Lynn PB, Jorge JM, et al. Impact of Organ-Preserving Strategies on Anorectal Function in Patients with Distal Rectal Cancer Following Neoadjuvant Chemoradiation. *Dis Colon Rectum*. Apr 2016;59(4):264-9.

PART I

PATIENT SELECTION AND FOLLOW-UP

CHAPTER 2

Selection of patients for organ preservation after chemoradiotherapy: MRI identifies poor responders who can go straight to surgery

Hester E. Haak, Monique Maas , Max J. Lahaye, Thierry N. Boellaard, Andrea Delli Pizzi, Casper Mihl, Dennis van der Zee, Cristina Fabris, Marit E. van der Sande, Jarno Melenhorst, Regina G. H. Beets-Tan, Geerard L. Beets, Doenja M. J Lambregts

Ann Surg Oncol. 2020 Aug;27(8):2732-2739

ABSTRACT

Objective

The aim of this study was to evaluate whether magnetic resonance imaging (MRI) can accurately identify poor responders after chemo radiotherapy (CRT) who will need to go straight to surgery, and to evaluate whether results are reproducible amongst radiologists with different levels of expertise.

Methods

Seven independent readers with different levels of expertise retrospectively evaluated the restaging MRIs (T2-weighted + diffusion-weighted imaging [T2W + DWI]) of 62 patients and categorized them as (1) poor responders – highly suspicious of tumor; (2) intermediate responders – tumor most likely; and (3) good – potential (near) complete responders. The reference standard was histopathology after surgery (or long-term follow-up in the case of a watch-and-wait program).

Results

Fourteen patients were complete responders and 48 had residual tumour. The median percentage of patients categorized as 'poor', 'intermediate' and 'good' responders by the 7 readers was 21% (range 11-37%), 50% (range 23-58%) and 29% (range 23-42%), respectively. The vast majority of poor responders had histopathologically confirmed residual tumour (73% ypT3-4), with a low rate (0-5%) of 'missed complete responders'. Of the 14 confirmed complete responders, a median percentage of 71% were categorized in the MR-good response and 29% in the MR-intermediate response group.

Conclusions

Radiologists of varying experience levels should be able to use MRI to identify the ±20% subgroup of poor responders who will definitely require surgical resection after CRT. This may facilitate a more selective use of endoscopy, particularly in general settings or in centers with limited access to endoscopy.

INTRODUCTION

In 2004, Habr Gama et al. introduced the concept of "watch-and-wait" (W&W) in rectal cancer, where patients with a clinical complete response (cCR) after neoadjuvant chemoradiotherapy (CRT) are deferred from surgery and instead closely monitored.[1] Since then, the W&W strategy has been successfully adopted by other clinical research groups[2-5] and W&W patients from 47 centers worldwide are now being registered in the International Watch and Wait Database (IWWD), the results of which, after the first 1000 inclusions, were recently published.[6]

One of the key issues in the W&W approach is how to best select the right candidates. Methods to assess response differ between published reports and centers [6-8] but data from the IWWD showed that endoscopy and MRI are the two tools most frequently used. Endoscopy is the most powerful tool to allow detailed assessment of the luminal response. MRI, particularly when combined with diffusion-weighted imaging (DWI), is a valuable adjunct to assess the lumen, and is of added benefit to diagnose any extraluminal findings, such as remaining positive lymph nodes, that may render W&W less feasible. In 64% of registered patients in the IWWD, a combination of both MRI and endoscopy was employed; it is generally acknowledged that this combination, together with clinical evaluation of the tumour, offers the best overall diagnostic performance to assess a complete (or near complete) response after chemoradiotherapy.[6, 9] Although both endoscopy and MRI are included in the selection process in most highly specialized centers that offer W&W as an alternative to resection, this cannot always be easily implemented in less specialized centers and with limited access to both modalities. The question is whether we can be more selective in the use of endoscopy. For example, it could be argued whether an endoscopy is necessary if MRI can accurately show that the patient has gross residual disease and needs to go straight for surgery. If this would be the case, it is critical that the multidisciplinary management team can rely on the findings of the radiologist and that the performance of the expert can be generalized.

Therefore, the primary goal of our study was to evaluate whether MRI can be used to accurately identify gross residual disease (poor response) after CRT, with a secondary goal of testing the reproducibility of MRI among radiologists with different levels of expertise.

METHODS

The retrospective use of imaging data for the purpose of this study was approved by the local Ethical Review Board and informed consent was waived.

Patient selection

The hospital's database (2011—2016) was searched for all non-metastatic, locally advanced, and/or distal rectal cancer patients who were diagnosed, staged with a standardized MRI protocol, and treated with long course neoadjuvant treatment at Maastricht University

Medical center. Inclusion criteria for this retrospective study consisted of (1) biopsy proven rectal adenocarcinoma; (2) neoadjuvant treatment consisting of long course CRT or short-course radiotherapy with a prolonged waiting interval of at least 6 weeks, (3) availability of a good-quality restaging MRI including a DWI sequence; and (4) availability of a valid standard of reference to establish the final response outcome, consisting of either histopathology after surgery (performed within 50 days following MRI) or a sustained cCR during long-term (> 2 year) follow-up in case of inclusion in a W&W program. Based on these inclusion criteria, a total of 62 eligible patients was identified.

Magnetic Resonance Imaging

All MRI examinations were performed on a 1.5T MR system (Philips Healthcare, Best, The Netherlands), according to protocols previously reported.[10, 11] In short, the protocol consisted of standard T2-weighted (T2W) turbo spin echo sequences in three directions (axial, sagittal, coronal; 3-5 mm slice thickness), an axial echo planar imaging (EPI) DWI sequence with the highest b-value being b1000 (slice thickness 5 mm), and corresponding Apparent Diffusion Coefficient (ADC) map calculated from the DWI sequence. The transverse T2W and DW axial images were angled in identical planes, perpendicular to the tumour axis, as identified on the sagittal T2W MRI. Since March 2014, patients routinely received a micro-enema prior to scanning (Microlax®; McNeil Healthcare, Ireland) to avoid susceptibility artefacts on DWI.[11] The micro-enema consisted of a 5ml solution, and was self-administered by patients approximately 15 minutes prior to MRI. No other bowel preparation or spasmolytic agents were used.

Image evaluation

MR images were read by seven independent readers with different levels of expertise (one resident reader, one abdominal radiologist working at a general non-academic center, two abdominal radiologists working at a general academic center, one abdominal radiologist working at a oncologic referral center, and two rectal MR experts working at a oncologic referral center). The readers were blinded to each other's results, the treatment following CRT (surgery or W&W), and the final response outcomes. Readers were asked to assess response on the restaging MR images, using a simplified three-category response system for which the readers were provided with a case report form (CRF) that was constructed for the purpose of this study. The CRF (illustrated in Figure 1) included imaging examples of all three response categories and was composed taking into account findings from previous publications on T2W MRI (including the MRI tumor regression grade [mrTRG]) and diffusion-weighted MRI response patterns.[12-14] Patients were categorized as:

(a) poor responders, i.e. patients with a high risk of residual tumour (in whom surgery will typically be required);
(b) intermediate responders, i.e. patients with an intermediate risk for residual tumour (in whom surgery will likely be required for the majority);
(c) good responders, i.e. patients who may have a near-complete or complete response.

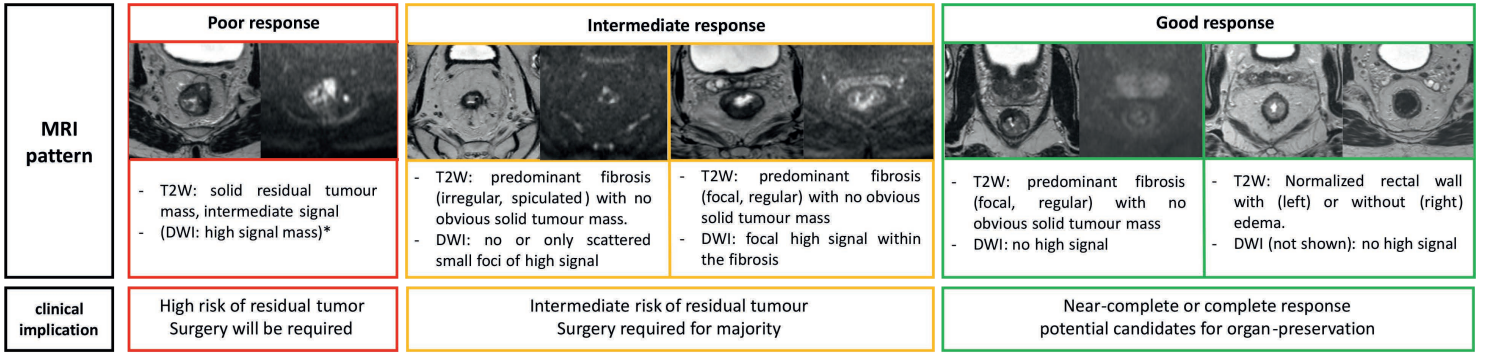


Figure 1 Three-category case report form to differentiate between poor, intermediate and good responders based on MRI. The response categories take into account the morphology on T2W MRI (in concordance with the mrTRG score, with poor response corresponding to mrTRG 4-5, intermediate response corresponding to mrTRG3, and good response to mrTRG 1-2 [13,14]) as well as previously reported DWI signal patterns.[12] * For the assessment of a poor response, with an obvious solid residual mass, availability of DWI is not mandatory to make the diagnosis. T2W T2-weighted imaging, DWI diffusion weighted imaging, MRI magnetic resonance imaging, mrTRG MRI tumor regression grade

In addition, the readers were asked to document any suspicious extraluminal findings, including pathologic lymph nodes (defined as any nodes ≥ 5 mm, according to recent guidelines published by the European Society of Gastrointestinal and Abdominal Radiology [15]), mesorectal tumour deposits, or the presence of gross extraluminal tumour extension (including extramural venous invasion).

Correlation with final response outcome

The final response outcome (complete response [CR] vs. non-CR/residual tumour) was defined based on histopathology (ypT0 vs. ypT1-4) in the surgically managed patients and on clinical follow-up in patients with a cCR who were included in a W&W program. In the latter group, a local regrowth-free follow-up period of at least 2 years was considered a surrogate endpoint for a CR (yT0N0)[6]. In the operated patients, the TRG[16] was also documented when available.

Data analysis

Statistical analyses were performed using IBM SPSS Statistics version 22.0 (IBM Corporation, Armonk, NY, USA). Descriptive statistics and contingency tables were constructed to compare the findings of the seven different readers and the patients' final response outcomes. Extraluminal findings on MRI were classified as true positive when positive nodes or deposits were confirmed at histopathology, or when the pathological resection specimen indicated a yT3-4 tumour to confirm extraluminal tumour extension on MRI. Interobserver agreement between individual readers was calculated using a weighted Kappa method with quadratic kappa weighting. Overall agreement between the seven readers was calculated with Kendall's coefficient of concordance.

RESULTS

Demographics

Demographics of the 62 study patients are shown in Table 1. Forty-one patients were male (66%) and median age was 67 years (range 45-83). In total, 14 (23%) patients were complete responders: 3 with ypT0 after surgery and 11 with a sustained ycT0 undergoing W&W, the latter with a median follow-up period of 49 months (range 35-66) at the time of writing. Forty-eight (77%) patients had a residual tumour after surgery (of whom 45 underwent immediate surgery and 3 had a regrowth within 3, 4 and 6 months, respectively, after initial inclusion in a W&W program). In the residual tumour group, 3 patients had ypT1 (6%), 15 ypT2 (31%), 27 ypT3 (57%), and 3 ypT4 disease (6%). The TRG was 2 in 13 (27%) patients, 3 in 15 (31%) patients, 4 in 14 (30%) patients and 5 in two (4%) patients; in 4 (8%) patients, the TRG was missing. Of the operated patients, 16 patients had N-positive disease.

Correlation between MR Response categories and final response outcome

Figure 2 shows the correlation of the MR scores of the seven different readers, with the final response outcome. The median percentage of patients categorized into the poor, intermediate and good response groups by the seven readers was 21% (range 11-37%), 50% (range 23-58%) and 29% (range 23-42%) respectively. When considering the total of 14 patients with a proven CR, the median percentage of these patients when categorized into the poor, intermediate and good response groups was 0%, 29%, and 71%, respectively, indicating that the majority were correctly classified as good responders. Apart from one patient with a CR who was misclassified in the poor response group by one of the seven readers, all patients categorized as MR-poor responders had confirmed residual tumour at histopathology, of whom the majority had advanced disease at histopathology (73% ypT3-4 tumours). The majority (76-100% for the seven different readers) of the MR-intermediate responders also had confirmed residual tumour, of which 58% still had ypT3-4 disease.

Extraluminal findings

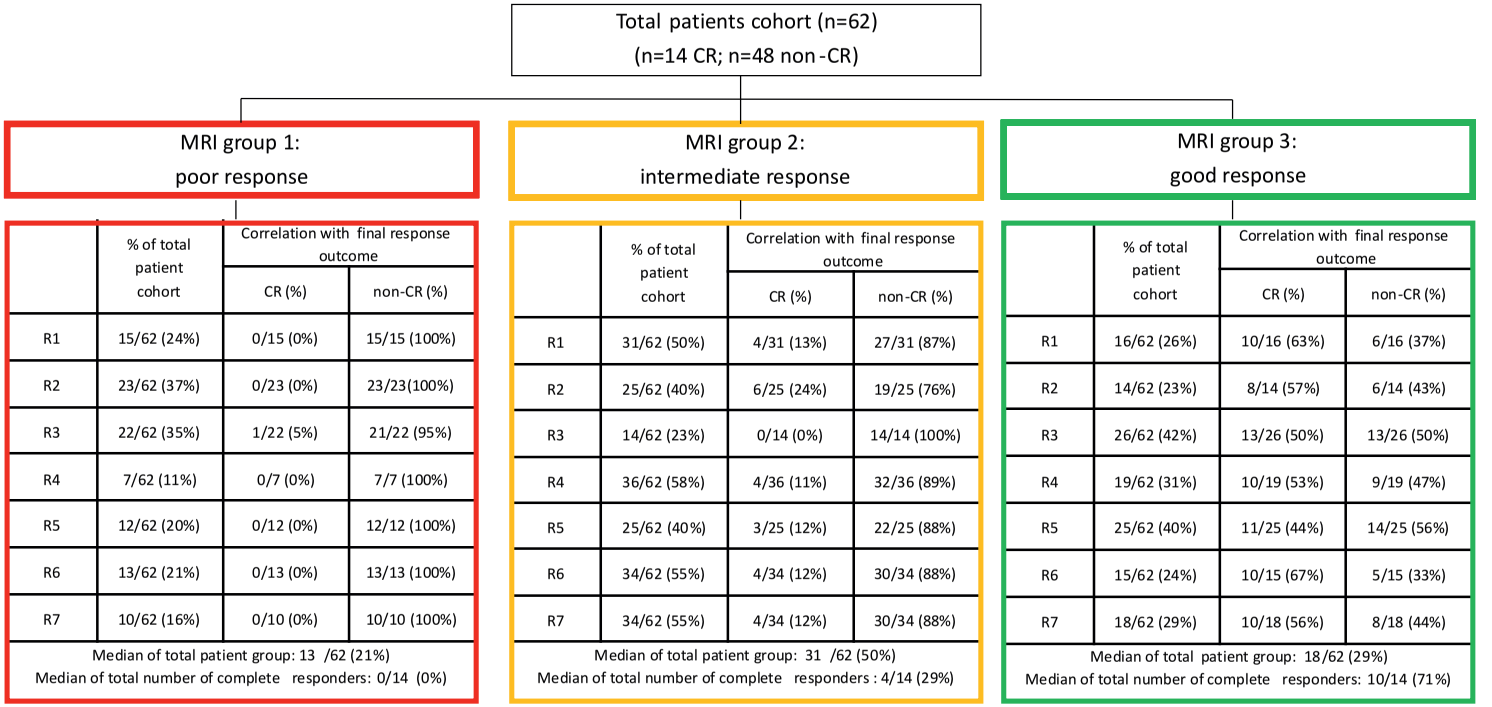
Table 2 describes the extraluminal findings as reported by the seven different readers. All patients (100% for all readers) who were scored as having extraluminal tumour extension had confirmed ypT3-4 residual disease at histopathology. Extraluminal tumour extension was only observed in the MR-poor responders (40-69%) and MR-intermediate responders (3-9%). None of the patients in the MR-good response group had any extraluminal tumour extension on MRI. In the good, intermediate and poor response groups, the seven readers identified positive nodes (or tumour deposits) in 0-19%, 13-36% and 8-50% of patients, respectively, which resulted in false positive rates ranging between 0-29% for the different readers, as illustrated in Table 2.

Interobserver agreement

Kendall's coefficient showed substantial overall agreement between the seven readers (W 0.65). Quadratic weighted kappa values between the different individual readers are listed in Table 3. Agreement between the most experienced readers (readers 1-3) was good (κ 0.64-0.68), while agreement between the remaining readers was moderate (κ 0.48-0.60), except for fair agreement between readers 2 and 5 (κ 0.38) and good agreement between readers 1 and 7 and readers 3 and 4 (κ 0.64 and κ 0.67, respectively).

Variables		Total [n=62]	
Age, years			
Median (range)		67 (45 - 83)	
Sex			
Male		41 (66)	
Female		21 (34)	
cT stage at primary staging			
1-2		8 (13)	
3		48 (77)	
4		6 (10)	
cN stage at primary staging			
0		13 (21)	
1		8 (13)	
2		41 (66)	
Neoadjuvant treatment			
5 x 5 Gy		7 (11)	
CRT		55 (89)	
Time between last Rtx and restaging MRI (days)			
Median (range)		56 (48 - 137)	
Time between restaging MRI and surgery (days)			
Median		16 (6 - 50)	
Final treatment			
W&W		14 (23)	
Immediate surgery		48 (77)	
Final response outcome			
CR		14 (23)	
	W&W	11	
	pCR	3	
Non-CR		48 (77)	
	Primary surgery	45	
			3ypT1
			13ypT2
			26ypT3
			3ypT4
	Delayed surgery (W&W with regrowth < 1 yr)	3	
			2ypT2
			1ypT3

Table 1 Patient demographics. Data are expressed as n (%) unless otherwise specified. CRT chemoradiation, W&W watch and wait, CR complete responders, pCR pathological complete response, MRI magnetic resonance imaging



	% of total patient cohort	Correlation with final response outcome	
		CR (%)	non-CR (%)
R1	15/62 (24%)	0/15 (0%)	15/15 (100%)
R2	23/62 (37%)	0/23 (0%)	23/23(100%)
R3	22/62 (35%)	1/22 (5%)	21/22 (95%)
R4	7/62 (11%)	0/7 (0%)	7/7 (100%)
R5	12/62 (20%)	0/12 (0%)	12/12 (100%)
R6	13/62 (21%)	0/13 (0%)	13/13 (100%)
R7	10/62 (16%)	0/10 (0%)	10/10 (100%)
Median of total patient group: 13 /62 (21%) Median of total number of complete responders: 0/14 (0%)			

	% of total patient cohort	Correlation with final response outcome	
		CR (%)	non-CR (%)
R1	31/62 (50%)	4/31 (13%)	27/31 (87%)
R2	25/62 (40%)	6/25 (24%)	19/25 (76%)
R3	14/62 (23%)	0/14 (0%)	14/14 (100%)
R4	36/62 (58%)	4/36 (11%)	32/36 (89%)
R5	25/62 (40%)	3/25 (12%)	22/25 (88%)
R6	34/62 (55%)	4/34 (12%)	30/34 (88%)
R7	34/62 (55%)	4/34 (12%)	30/34 (88%)
Median of total patient group: 31 /62 (50%) Median of total number of complete responders : 4/14 (29%)			

	% of total patient cohort	Correlation with final response outcome	
		CR (%)	non-CR (%)
R1	16/62 (26%)	10/16 (63%)	6/16 (37%)
R2	14/62 (23%)	8/14 (57%)	6/14 (43%)
R3	26/62 (42%)	13/26 (50%)	13/26 (50%)
R4	19/62 (31%)	10/19 (53%)	9/19 (47%)
R5	25/62 (40%)	11/25 (44%)	14/25 (56%)
R6	15/62 (24%)	10/15 (67%)	5/15 (33%)
R7	18/62 (29%)	10/18 (56%)	8/18 (44%)
Median of total patient group: 18/62 (29%) Median of total number of complete responders: 10/14 (71%)			

Figure 2 Response categorization results for the seven respective readers versus the final response outcome. Rx reader x, CR complete responders

Extraluminal findings	Patients classified in the MR- poor response group		Patients classified in the MR-intermediate response group		Patients classified in the MR-good response group	
Nodes or deposits	**Total (%)**	**FP (%) ***	**Total (%)**	**FP (%) ***	**Total (%)**	**FP (%) ***
R1	2/15 (13%)	0 (0%)	4/31 (13%)	1 (3%)	3/16 (19%)	2 (13%)
R2	4/23 (17%)	1 (4%)	6/25 (24%)	1 (4%)	2/14 (14%)	1 (7%)
R3	9/22 (41%)	2 (9%)	4/14 (29%)	0 (0%)	1/26 (4%)	0 (0%)
R4	2/7 (29%)	2 (29%)	13/36 (36%)	4 (11%)	2/19 (11%)	1 (5%)
R5	6/12 (50%)	1 (8%)	8/25 (32%)	5 (20%)	2/25 (8%)	0 (0%)
R6	1/13 (8%)	1 (8%)	10/34 (29%)	3 (9%)	0/15 (0%)	N/A
R7	4/10 (40%)	0 (0%)	7/34 (21%)	1 (3%)	3/18 (17%)	2 (1%)
Extraluminal tumour extension	**Total (%)**	**FP (%) ****	**Total (%)**	**FP (%) ****	**Total (%)**	**FP (%) ****
R1	6/15 (40%)	0 (0%)	1/31 (3%)	0 (0%)	0/16 (0%)	N/A
R2	11/23 (48%)	0 (0%)	2/25 (8%)	0 (0%)	0/14 (0%)	N/A
R3	11/22 (50%)	0 (0%)	1/14 (7%)	0 (0%)	0/26 (0%)	N/A
R4	4/7 (57%)	0 (0%)	1/36 (3%)	0 (0%)	0/19 (0%)	N/A
R5	5/12 (42%)	0 (0%)	2/25 (8%)	0 (0%)	0/25 (0%)	N/A
R6	9/13 (69%)	0 (0%)	1/34 (3%)	0 (0%)	0/15 (0%)	N/A
R7	4/10 (40%)	0 (0%)	3/34 (9%)	0 (0%)	0/18 (0%)	N/A

Table 2 Extraluminal findings. * yN1-N2 detected in pathological specimens, ** yT3-T4 detected in pathological specimens. FP false positives, Rx reader x, NA not applicable

	R1	R2	R3	R4	R5	R6	R7
R1	NA	0.65	0.68	0.56	0.48	0.61	0.64
R2			0.64	0.56	0.38	0.53	0.57
R3				0.67	0.53	0.60	0.55
R4					0.54	0.59	0.53
R5						0.49	0.53
R6							0.57
R7							NA

Table 3 Interobserver agreement between readers. R1 reader 1, rectal MRI expert working in an oncologic referral center, R2 reader 2, rectal MRI expert working in an oncologic referral center, R3 reader 3, abdominal radiologist working in an oncologic referral center, R4 reader 4, abdominal radiologist working in an academic center, R5 reader 5, abdominal radiologist working in an academic center, R6 reader 6, abdominal radiologist working in non-academic center, R7 reader 7, radiologist trainee with no specific MRI expertise, NA not applicable, MRI magnetic resonance imaging

DISCUSSION

This study has shown that although agreement between individual readers was not always perfect, radiologists with varying levels of expertise in interpreting rectal cancer MRIs were able to correctly identify the ±20% of poor responders who will definitely require surgery and typically present with substantial (ypT3-4) residual disease at histopathology.

Interpretation of MRIs after CRT is well-known to be hampered by difficulties in discerning fibrosis from residual disease. Different MR interpretation and classification systems have been suggested focusing on specific morphological T2W-MRI patterns (including the mrTRG) and/or DWI signal patterns to assess response after CRT.[12-14, 17] However, these systems require a certain level of expertise, and, in particular, for DWI there are some known pitfalls that may lead to misinterpretations.[10] The results as reported for expert readers in published reports may therefore be less reproducible in less experienced hands, and may be difficult to translate to general everyday practice.

Our study shows that when using a simplified three-category response evaluation system to make a more approximate estimate of the risk of residual disease (Figure 1), all readers, regardless of the level of expertise, were able to identify, on MRI, the group of poor responders with gross residual disease. Moreover, of the 14 confirmed complete responders, the majority (71%) were correctly categorized into the MR-good response group and the remaining 29% were categorized into the MR-intermediate response group. Together, these patients thus represent the largest subgroup of ±80% of patients who benefit most from detailed response evaluation with endoscopy combined with MRI, to make a fully informed decision between TME, local excision in case of a small residual tumour lesion, or W&W in case of a confirmed cCR. Although some researchers have reported that patients with a CR may still show some mucosal abnormalities on endoscopy, and that false negative and

positive biopsies may occur[18, 19], endoscopy is generally acknowledged as an invaluable tool to assess luminal response after CRT in rectal cancer[9, 20]. In the setting of organ preservation, the combined use of clinical evaluation, MRI and endoscopy thus remains the preferred and most accurate selection method. In a setting where there is less access to both selection modalities, one could be more selective with the use of endoscopy and refer patients straight for surgery based on MRI only if the MRI shows gross residual disease after CRT.

In addition to luminal response assessment, MRI is particularly valuable for identifying extraluminal tumours and remaining mesorectal nodes/deposits, which could be a contraindication for W&W. In the current study, all patients with MR-detected extraluminal tumour extension had confirmed ypT3-T4 tumour according to histopathology. Our results regarding the detection of remaining vital lymph node metastases and tumour deposits were unfortunately not so good and a variable number of false positive findings occurred, which is in line with the known inaccuracies of MRI for nodal staging.[21, 22] Nevertheless, the number of false positive findings in the MR-good response group was low, ranging between only 0-13% for the seven different readers.

This study has several limitations, in addition to its relatively small-sized cohort and retrospective nature. First, not all patients had histopathological confirmation. However, the patients with a sustained cCR all had a follow up of > 2 years (range 35-66 months), which will generally be considered a good surrogate endpoint of a complete remission as most regrowths are known to occur in the first 2 years.[6] Second, our study design was based on a clinical scenario assuming routine use of MRI as a first line response tool. One could argue that, depending on local policy and availability of the respective modalities, an alternative strategy applying endoscopy as a first-line tool with more selective use of MRI could be just as effective; however, exploring this alternative strategy was outside the scope of this retrospective study. Finally, although the readers in our study had varying levels of expertise, the majority were relatively experienced abdominal readers with at least an affinity for reading rectal MRIs. Prospective and large-scale validation will therefore be required to further validate our findings in more general clinical settings.

CONCLUSIONS

Our study suggests that regardless of their level of expertise, radiologists should be able to accurately identify, on MRI, the ±20% subgroup of patients with gross residual disease who can go straight for surgery and who would benefit less from further endoscopic assessment. We support previous evidence that for the remaining majority of patients, a combined use of clinical evaluation, MRI, and endoscopy remains the preferred response evaluation method when aiming to select patients for organ preservation (W&W). Once validated prospectively, such an approach could allow more selective use of diagnostic tools, thereby facilitating the implementation of W&W in busy everyday practice.

REFERENCES

1. Habr-Gama A, Perez RO, Nadalin W, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg*. Oct 2004;240(4):711-7; discussion 717-8.
2. Appelt AL, Pløen J, Harling H, et al. High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. *The Lancet Oncology*. 2015/08/01/ 2015;16(8):919-927.
3. Renehan AG, Malcomson L, Emsley R, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *The Lancet Oncology*. 2016;17(2):174-183.
4. Martens MH, Maas M, Heijnen LA, et al. Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. *J Natl Cancer Inst*. Dec 2016;108(12).
5. Smith JJ, Strombom P, Chow OS, et al. Assessment of a Watch-and-Wait Strategy for Rectal Cancer in Patients With a Complete Response After Neoadjuvant Therapy. *JAMA oncology*. Jan 10 2019:e185896.
6. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
7. Dossa F CTR, Acuna S.A, Baxter N.N. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete respons following neoadjuvant chemoradiation: a systematic review and meta-analysis. *Lancet Gastroenterology heptology*. 2017;(2):501-13.
8. Stijns RCH, Tromp MR, Hugen N, de Wilt JHW. Advances in organ preserving strategies in rectal cancer patients. *Eur J Surg Oncol*. Feb 2018;44(2):209-219.
9. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol*. Nov 2015;22(12):3873-80.
10. Lambregts DMJ, Van Heeswijk MM, Delli Pizzi A, et al. Diffusion-weighted MRI to assess response to chemoradiotherapy in rectal cancer: main interpretation pitfalls and their use for teaching. *Eur Radiol*. 2017;27 (10):4445-4454.
11. van Griethuysen JJM, Bus EM, Hauptmann M, et al. Gas-induced susceptibility artefacts on diffusion-weighted MRI of the rectum at 1.5T - Effect of applying a micro-enema to improve image quality. *Eur J Radiol*. Feb 2018;99:131-137.
12. Lambregts DMJ, Delli Pizzi A, Lahaye MJ, et al. A Pattern-Based Approach Combining Tumor Morphology on MRI With Distinct Signal Patterns on Diffusion-Weighted Imaging to Assess Response of Rectal Tumors After Chemoradiotherapy. *Dis Colon Rectum*. Mar 2018;61(3):328-337.
13. Siddiqui MR, Gormly KL, Bhoday J, et al. Interobserver agreement of radiologists assessing the response of rectal cancers to preoperative chemoradiation using the MRI tumour regression grading (mrTRG). *Clin Radiol*. Sep 2016;71(9):854-62.
14. Patel UB, Taylor F, Blomqvist L, et al. Magnetic resonance imaging-detected tumor response for locally advanced rectal cancer predicts survival outcomes: MERCURY experience. *J Clin Oncol*. Oct 1 2011;29(28):3753-60.

15. Beets-Tan RGH, Lambregts DMJ, Maas M, et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *Eur Radiol*. Apr 2018;28(4):1465-1475.
16. Mandard AM, Dalibard F, Mandard JC, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer*. Jun 1 1994;73(11):2680-6.
17. Santiago I, Barata M, Figueiredo N, et al. The split scar sign as an indicator of sustained complete response after neoadjuvant therapy in rectal cancer. *Eur Radiol*. 2010;30(1):224-238.
18. van der Sande ME, Beets GL, Hupkens BJ, et al. Response assessment after (chemo)radiotherapy for rectal cancer: Why are we missing complete responses with MRI and endoscopy? *Eur J Surg Oncol*. 2019;45(6):1011-1017.
19. Smith FM, Wiland H, Mace A, Pai RK, Kalady MF. Clinical criteria underestimate complete pathological response in rectal cancer treated with neoadjuvant chemoradiotherapy. *Dis Colon Rectum*. Mar 2014;57(3):311-315.
20. Smith FM, Chang KH, Sheahan K, Hyland J, O'Connell PR, Winter DC. The surgical significance of residual mucosal abnormalities in rectal cancer following neoadjuvant chemoradiotherapy. *Br J Surg*. Jul 2012;99(7):993-1001.
21. Bipat S, Glas AS, Slors FJ, Zwinderman AH, Bossuyt PM, Stoker J. Rectal cancer: local staging and assessment of lymph node involvement with endoluminal US, CT, and MR imaging--a meta-analysis. *Radiology*. Sep 2004;232(3):773-83.
22. van der Paardt MP, Zagers MB, Beets-Tan RG, Stoker J, Bipat S. Patients who undergo preoperative chemoradiotherapy for locally advanced rectal cancer restaged by using diagnostic MR imaging: a systematic review and meta-analysis. *Radiology*. Oct 2013;269(1):101-12.

CHAPTER 3

Prevalence of nodal involvement in rectal cancer after chemoradiotherapy

Hester E. Haak, Geerard L. Beets, Koen Peeters, Patty J. Nelemans, Vincenzo Valentini, Claus Rödel, Li-Jen Kuo, Felipe A. Calvo, Julio Garcia-Aguilar, Rob Glynne-Jones, Salvatore Pucciarelli, Javier Suarez, George Theodoropoulos, Sebastiano Biondo, Doenja M. J. Lambregts, Regina G. H. Beets-Tan, Monique Maas

Br J Surg. 2021 Jul 8:znab194.

ABSTRACT

Background

The purpose of this study was to investigate the prevalence of ypN+ status according to ypT category in patients with locally advanced rectal cancer treated with chemoradiotherapy and total mesorectal excision, and to assess the impact of ypN+ on disease recurrence and survival by pooled analysis of individual-patient data.

Methods

Individual-patient data from 10 studies of chemoradiotherapy for rectal cancer were included. Pooled rates of ypN+ disease were calculated with 95 per cent confidence interval for each ypT category. Kaplan-Meier and Cox regression analyses were undertaken to assess influence of ypN status on 5-year disease-free survival (DFS) and overall survival (OS).

Results

Data on 1898 patients were included for the study. Median follow-up was 50 (range 0-219) months. The pooled rate of ypN+ disease was 7 per cent for ypT0, 12 per cent for ypT1, 17 per cent for ypT2, 40 per cent for ypT3 and 46 per cent for ypT4. Patients with ypN+ disease had lower 5-year DFS and OS (46.2 and 63.4 per cent respectively) than patients with ypN0 tumours (74.5 and 83.2 per cent respectively) ($P<0.001$). Cox regression analyses showed ypN+ status to be an independent predictor of recurrence and death.

Conclusion

Risk of nodal metastases (ypN+) after chemoradiotherapy increases with advancing ypT category and needs to be considered if an organ-preserving strategy is contemplated.

INTRODUCTION

Total mesorectal excision (TME) and neoadjuvant (chemo)radiotherapy have improved rectal cancer treatment[1,2] by reducing local failure rates. Neoadjuvant therapy may also facilitate organ-preservation strategies, whereby adequate local control may be achieved without the morbidity and quality-of-life implications associated with surgery.[3-5] For patients with a good response but a small residual lesion, some believe that local excision of the residual disease is appropriate, provided that regional lymph nodes have been sterilized with chemoradiotherapy.[6, 7] In primary early rectal cancer, the baseline tumour characteristics (T-category) can be used to estimate the risk of lymph node metastases (N status). This helps both the selection of patients for primary treatment by local excision and in the decision whether or not to perform a completion TME after local excision.[8-14] The same strategy could be used for patients with a small residual tumour after chemoradiotherapy, but fewer data are available regarding the prevalence of ypN+ among small residual lesions in patients with a locally advanced tumour at baseline. Overall, ypN+ rates reported in the literature vary from 0 to 11 per cent for ypT1 disease, 8 to 29 per cent for ypT2 disease and 37 to 40 per cent for ypT3 disease.[15-17] To gain more insight into the risk of ypN+ status in locally advanced rectal cancer, this study investigated the prevalence of ypN+ according to ypT category in patients with locally advanced rectal cancer treated with chemoradiotherapy and TME, by a pooled analysis of individual-patient data.

METHODS

Patient data were selected from a data set that was used for a pooled meta-analysis with individual-patient data examining the prognostic significance of a complete response after chemoradiotherapy for patients with locally advanced cancer.[18] As the study contained data from previously published studies, no ethics approval or patient consent was needed. In total, 14 studies were included in the original study by Maas and colleagues[18], of which 10[19-28] could be included in the present analysis. One study was excluded because only patients with ypN0 were included, two studies were excluded because of missing data on ypT categories (other than ypT0 versus ypT+) or missing information on receipt of adjuvant chemotherapy, and the author of another study declined participation for this analysis. The data from previous studies were combined into a single data set. The data comprised patient characteristics, baseline staging data, treatment details, histological data, and follow-up details.

Statistical analyses

The frequency of ypN+ status according to ypT category was calculated for each study, and pooled for all studies with 95 per cent confidence intervals by use of a random-effects model. To stabilize the variance of the proportions from individual studies, Freeman-Tukey arcsine square root transformation of the proportion with ypN+ status was used.[29] The

transformed proportions were pooled using a DerSimonian and Laird random-effects model to account for heterogeneity among studies.[30] Heterogeneity was quantified by the I^2 index and Cochran's Q test.[31] For comparison of 5-year cumulative probability of local and distant recurrence, as well as disease-free survival (DFS) and overall survival (OS) between patients with ypN+ and ypN0 status, Kaplan-Meier analysis and Cox proportional hazards models stratified by study were used. For these time-to-event analyses, follow-up started on the day of surgery and ended on the day of disease relapse or death or day of last follow-up. Patients were censored if, by the end of the follow-up period, they had not developed the outcome of interest or if they were lost to follow-up. The log rank test was used to compare Kaplan-Meier curves. The Cox proportional hazards assumption was tested on the basis of Schoenfeld residuals after fitting a model and by visual inspection of log minus log plots. The proportional hazards assumption is not violated if the proportionality test is not significant and the plots show that the survival curves for groups being compared run parallel to each other. $P \leq 0.050$ was considered statistically significant. Analyses were performed using StatsDirect® software (StatsDirect, Altrincham, UK).

RESULTS

Patient and treatment characteristics for each study are shown in Table 1 and Table S1. The imaging technique used for clinical staging varied between studies; it mainly consisted of endorectal ultrasonography and CT, with additional MRI in some studies. A total of 2026 patients were included in the data sets of the original 10 selected studies, of whom 128 were excluded owing to unknown ypT or ypN category. Therefore, 1898 patients were included in the present analyses. Survival data were available for 1856 patients. All studies used external beam radiotherapy in doses ranging from 45 to 50.4 Gy in 25-28 fractions. The interval between chemoradiotherapy and surgery was most commonly 6-8 weeks. Chemotherapy using 5-fluorouracil was administered as a radiosensitizer in the majority of the patients. Most patients also received adjuvant chemotherapy (5-FU based); the type of adjuvant therapy was unknown for two studies.

Reference	Population	No. of patients	Type of neoadjuvant treatment	Interval between CRT and surgery (weeks)	Type of adjuvant chemotherapy	Type of study	Clinical staging modality
Valentini[22]	LARC, extraperitoneal, T3-T4 or N+	474	External RT or IORT, 5-FU + mitomycin C/ cisplatin	6-8	5-FU	Prospective	EUS + CT
Rödel[28]	Stage II-III	348	External RT, 5-FU	6	5-FU	Prospective (arm of RCT)	EUS + CT
Kuo[26]	T3-T4 N+M0	242	External RT, 5-FU, mitomycin C	6-8	n.r.	Retrospective	MRI
García-Aguilar[19]	Stage II-III	154	External RT, 5-FU	6	5-FU and leucovorin	Retrospective	EUS + CT
Glynne-Jones[20]	T3-T4	147	External RT, 5-FU	6-12	n.r.	Prospective	EUS + CT + MRI
Suárez[27]	LARC	119	External RT, 5-FU	6	n.r.	Retrospective	CT
Calvo[21]	T3-4 N+	117	External RT/IORT, 5-FU or tegafur	4-6	5-FU and leucovorin	Prospective	EUS + CT
Pucciarelli[23]	T3-4 N+M0	106	External RT, 5-FU + leucovorin/ carboplatin/ oxaliplatin	6-8	5-FU and leucovorin	Retrospective	EUS + CT
Biondo[25]	T3-4 LARC	103	External RT, 5-FU	6-8	5-FU and leucovorin	Prospective	CT
Theodoropoulos[24]	All	88	External RT, 5-FU +/- leucovorin	6	n.r.	Retrospective	EUS + CT

Table 1 Characteristics from the included studies. Adapted from M. Maas et al.[18] CRT, chemoradiotherapy; LARC, locally advanced rectal cancer; N+, clinically node-positive; RT, radiotherapy; IORT, intra-operative radiotherapy; 5-FU, 5-fluorouracil; EUS, endorectal ultrasound; n.r., not reported.

Of all 1795 patients with available data on cT category, 1708 (95.1 per cent) were diagnosed with cT3-4 disease before neoadjuvant treatment. Data on cN status were available for 1802 patients, of whom 1080 (59.9 per cent) had cN+ disease, whereas only 26.2% had ypN+ disease at histological examination of the resection specimen. Median follow-up was 50 (range 0-219) months.

The pooled rate of ypN+ disease was 7 (95 per cent c.i. 3 to 12) per cent for ypT0 (I^2 = 56 per cent; P = 0.015), 12 (4 to 24) per cent for ypT1 (I^2 = 53 per cent; P = 0.025), 17 (12 to 23) per cent for ypT2 (I^2 = 62 per cent; P = 0.005), 40 (36 to 44) per cent for ypT3 (I^2 = 32 per cent; P = 0.154) and 46 (34 to 57) per cent for ypT4 (I^2 = 0 per cent; P= 0.586) (Figure 1). Table 2 provides an overview of the proportion of patients with (y)pN+ disease according to (y)pT category after chemoradiotherapy in the present study, compared with rates reported in the literature for patients who did not receive neoadjuvant treatment.

Long-term outcome

Patients with ypN+ disease had a lower DFS and OS rates at 5 years than patients with ypN0 disease (Figure 2). Patients with cN+ tumours before chemoradiotherapy who had ypN0 status after chemoradiotherapy had similar 5-year DFS to patients who had cN0 lesions at primary staging and ypN0 after chemoradiotherapy: 74.8 (95 per cent c.i. 72 to 78) and 73.7 (70 to 78) per cent respectively. cN status had limited accuracy, reflected by the large number of patients staged as cN0 who had ypN+ disease after TME (156 of 722, 21.6 per cent). In addition, cN had only moderate predictive value for long-term DFS (hazard ratio (HR) 1.03, 95 per cent c.i. 0.84 to 1.28) and OS (HR 1.20, 0.94 to 1.54).

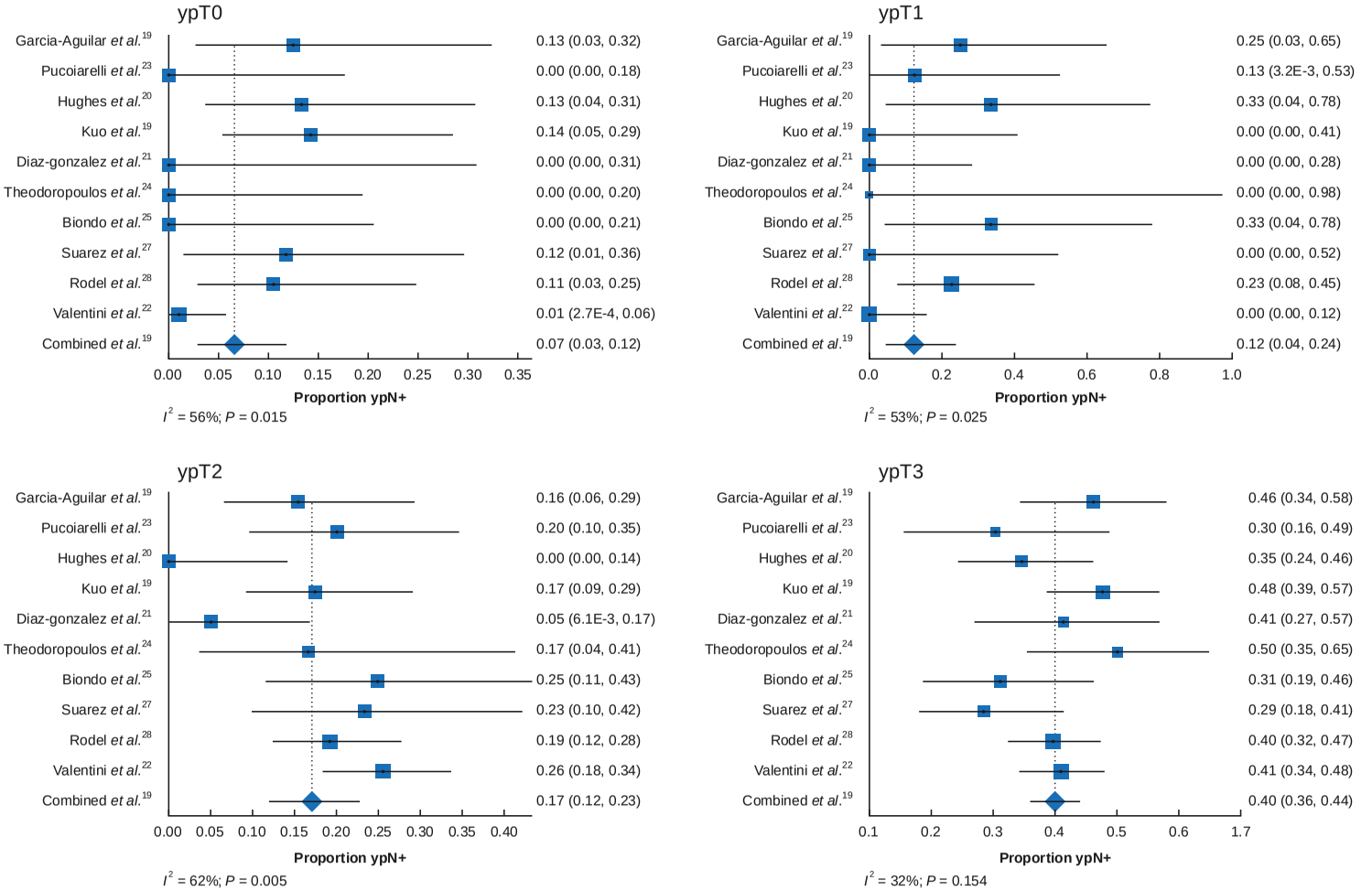


Figure 1 Pooled proportions of ypN+ disease by ypT category. **a** ypT0, **b** ypT1, **c** ypT2, **d** ypT3, and **e** ypT4. A random-effects model was used for meta-analysis. Proportions are shown with 95 per cent intervals. Heterogeneity is indicated by the I^2 value.

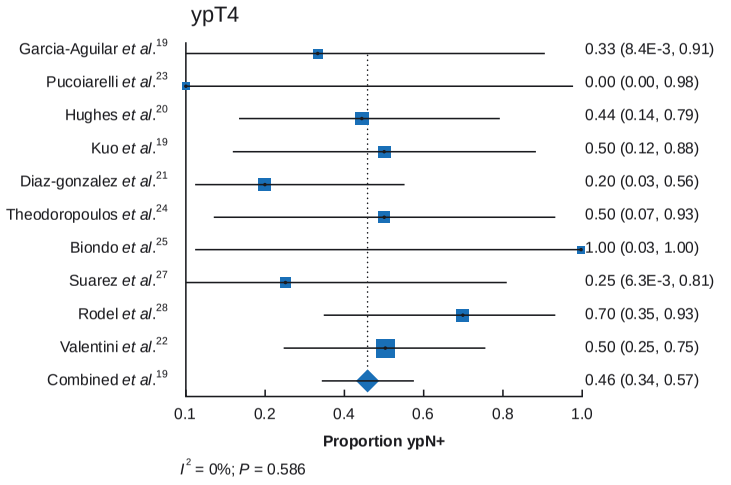


Figure 1 Continued.

	ypN+ rate (%)	
	After chemoradiotherapy (present study)	**Without neoadjuvant treatment (published studies)**
(y)pT0	7	
(y)pT1	12	6-14[8-14]
(y)pT2	17	17-23[8-14]
(y)pT3	40	49-66[8, 13]
(y)pT4	46	50-79[8, 13]

Table 2 Proportion of patients with positive lymph nodes according to (y)pT category after chemoradiation in the present study compared with results reported in the literature for patients not treated with neoadjuvant chemoradiotherapy.

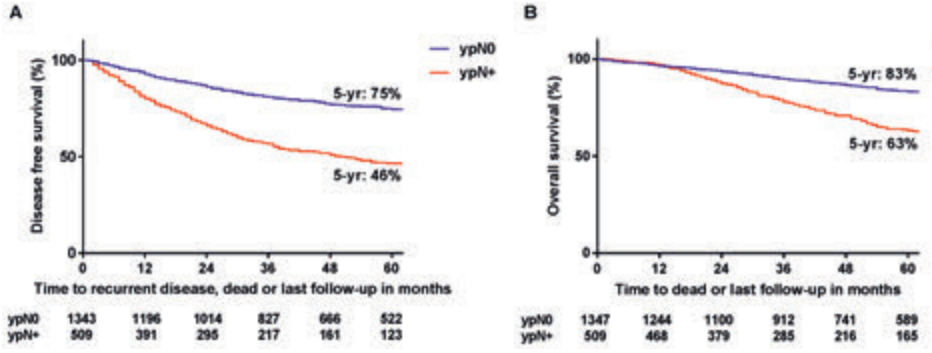


Figure 2 Survival curves by ypN status for the total patient group **a** Disease-free survival and **b** overall survival. **a,b** P<0.001 (log rank test).

In the subgroup of patients with ypT0-2 disease, there was a difference in 5-year DFS between ypN+ and ypN0 groups: 65.0 (57 to 74) and 81.3 (78 to 84) per cent respectively (P<0.001). Five-year OS rates also differed: 81.1 (73 to 88) versus 87.5 (85 to 90) per cent (P=0.005). Additional survival analyses according to ypN status separated by ypT category are described in Figure S1.

In a multivariable Cox regression model, stratified by centre (including sex, age, cT, cN, distance from anal verge, type of surgery, ypT and chemotherapy as independent variables), ypN+ status was a predictor of recurrence and death, with HRs of 2.45 (1.70 to 3.54) and 2.05 (1.28 to 3.29) for DFS and OS respectively in the subgroup of patients with ypT0-2 (Table 3), but also in the total patient group (Table 4).

		Hazard ratio	
		Disease-free survival	**Overall survival**
Sex			
	M	1.00 (reference)	1.00 (reference)
	F	0.84 (0.61, 1.17)	0.73 (0.49, 1.11)
Age (per year)		1.00 (0.99, 1.01)	1.00 (0.99, 1.02)
Clinical tumour category at baseline			
	cT1	0.54 (0.24, 1.24)	1.04 (0.34, 3.22)
	cT2	0.98 (0.51, 1.86)	0.60 (0.23, 1.58)
	cT3	1.00 (reference)	1.00 (reference)
	cT4	1.88 (1.20, 2.97)	1.58 (0.92, 2.74)
Clinical node category at baseline			
	cN0	1.00 (reference)	1.00 (reference)
	cN+	0.94 (0.67, 1.35)	1.14 (0.76, 1.73)
Distance from anal verge (cm)			
	≤5	1.00 (reference)	1.00 (reference)
	>5	1.09 (0.78, 1.55)	1.40 (0.93, 2.13)
Type of surgery			
	LAR	1.00 (reference)	1.00 (reference)
	APR	1.48 (1.00, 2.20)	1.81 (1.15, 2.89)
	Other	1.55 (0.73, 3.28)	2.21 (0.98, 5.04)
Pathological T category			
	pT0	1.00 (reference)	1.00 (reference)
	pT1	0.75 (0.42, 1.36)	0.77 (0.40, 1.48)
	pT2	1.10 (0.77, 1.58)	0.84 (0.56, 1.28)
Pathological N category			
	pN0	1.00 (reference)	1.00 (reference)
	pN+	2.45 (1.70, 3.54)	2.05 (1.28, 3.29)
Adjuvant chemotherapy			
	No	1.00 (reference)	1.00 (reference)
	Yes	0.64 (0.44, 0.96)	0.49 (0.30, 0.81)

Table 3 Adjusted hazard ratios from multivariable Cox proportional hazards models for patients with ypT0–2 disease stratified by data set. Values in parentheses are 95 per cent confidence intervals; LAR, low anterior resection; APR, abdominal perineal resection. A hazard ratio below 1 indicates a lower probability of an unfavourable event.

		Hazard ratio	
		Disease-free survival	Overall survival
Sex			
	M	1.00 (reference)	1.00 (reference)
	F	0.84 (0.70, 1.01)	0.81 (0.65, 1.02)
Age (per year)		0.99 (0.98, 1.00)	0.99 (0.99, 1.01)
Clinical tumour category at baseline			
	cT1	0.75 (0.43, 1.34)	1.27 (0.57, 2.84)
	cT2	0.99 (0.60, 1.63)	0.63 (0.30, 1.31)
	cT3	1.00 (reference)	1.00 (reference)
	cT4	1.33 (1.04, 1.72)	1.23 (0.92, 1.67)
Clinical node category at baseline			
	cN0	1.00 (reference)	1.00 (reference)
	cN+	1.03 (0.84, 1.28)	1.20 (0.94, 1.54)
Distance from anal verge (cm)			
	≤5	1.00 (reference)	1.00 (reference)
	>5	1.03 (0.85, 1.27)	1.16 (0.92, 1.48)
Type of surgery			
	LAR	1.00 (reference)	1.00 (reference)
	APR	1.52 (1.23, 1.90)	1.65 (1.27, 2.15)
	Other	1.23 (0.79, 1.94)	1.50 (0.89, 2.55)
Pathological T category			
	pT0	1.00 (reference)	1.00 (reference)
	pT1	0.85 (0.48, 1.51)	0.82 (0.44, 1.57)
	pT2	1.15 (0.82, 1.63)	0.85 (0.57, 1.28)
	pT3	2.01 (1.46, 2.77)	1.62 (1.13, 2.33)
	pT4	2.89 (1.77, 4.74)	2.37 (1.37, 4.11)
Pathological N category			
	pN0	1.00 (reference)	1.00 (reference)
	pN+	2.26 (1.87, 2.74)	2.08 (1.66, 2.62)
Adjuvant chemotherapy			
	No	1.00 (reference)	1.00 (reference)
	Yes	0.61 (0.49, 0.76)	0.51 (0.40, 0.68)

Table 4 Adjusted hazard ratios from multivariable Cox proportional hazards models for the total cohort stratified by data set. Values in parentheses are 95 per cent confidence intervals; LAR, low anterior resection; APR, abdominal perineal resection. A hazard ratio below 1 indicates a lower probability of an unfavourable event.

DISCUSSION

This study has shown that the pooled prevalence of lymph node metastases after neoadjuvant chemoradiotherapy in patients with locally advanced rectal cancer increases with increasing depth of residual tumour, and is in the same range as that for non-irradiated tumours. With a tumour complete response (ypT0) there is still a 7 per cent risk of lymph node metastases. In this setting, the presence of lymph node metastases is a strong predictor of poor long-term outcome, as for non-irradiated tumours.

The findings of this study are in accordance with previous reports. Generally, rates of lymph node metastases in patients with ypT0 disease are below 10 per cent in most studies[15, 32, 33]. In ypT2 tumours, lymph node metastases have been reported in up to 29 percent of patients[15, 16], which is higher than the 17 per cent in the present study. However, the GRECCAR 2 trial[17] reported a much lower incidence of nodal involvement of 8 per cent, which may be explained by differences in the study population as the GRECCAR 2 trial included patients with smaller tumours (less than 4 cm) with cT2–3 N0–1 stage, with at most limited nodal disease at diagnosis. The present study included more locally advanced tumours at diagnosis.

A focus on the prevalence of lymph node metastases is particularly relevant when organ preservation is being contemplated. With all organ-preserving strategies (including local scar excision) the regional lymph nodes are left in situ and are a potential source of recurrence. Although it is often stated that the risk of leaving involved nodes behind is small for ypT0–1 tumours and too high for ypT2 tumours, the differences were not that marked in the present study (7, 12, and 17 per cent for ypT0, ypT1, and ypT2 respectively). The prevalence of 40 per cent for ypT3 tumours was substantially higher. Whether or not to consider organ preservation or to undertake TME is reliant on a risk–benefit assessment that should include information from baseline and post-treatment staging, histology if local excision was performed, and also patient preference and co-morbidity.

It is also interesting to note that in a pooled analysis of 880 watch and wait patients with a clinical complete response only 11 patients had a nodal regrowth[3]. This is much lower than would be expected from the present findings. There are a number of possible reasons for this. Not all lymph node metastases detected by the pathologist in the TME specimen 6–8 weeks after irradiation may represent viable tumour, and the longer interval between restaging and the decision to watch and wait may allow further regression.[34] Residual macrometastases in nodes are associated with a poor prognosis. However, small residual micrometastases found in the nodes at histopathology 6–8 weeks after chemoradiotherapy might regress if a longer interval is applied, and may not be of clinical significance (62 per cent ypN0 within 4–8 weeks versus 73 per cent ypN0 within 8–12 weeks)[35, 36]. ypT category is also a crude measure of response to chemoradiotherapy that does not correlate directly with tumour volume. Patients who have an apparently (near) complete response at restaging (MRI and endoscopy) but actually have a small ypT2 remnant that becomes obvious with follow-up

could have a lower proportion of lymph node metastases than patients with a moderate response and a large remaining ypT2 tumour. Finally, although still controversial in early disease[37], MRI has improved local staging, so patients with obvious lymph node metastases on imaging are not selected for organ preservation and undergo formal TME, which reduces the risk of nodal regrowth. Regardless of the real prevalence of lymph node metastases in different organ preservation strategies, follow-up with serial MR imaging is essential to identify these patients as early as possible and to perform a delayed TME. In addition to ypT category, there are other histological parameters by which to identify patients at a higher risk of lymph node metastases who are less suitable for organ-preserving treatment, such as lymphatic or vascular invasion and differentiation grade.[38, 39] As differentiation grade and other histopathological factors of the tumour were poorly recorded in this pooled data set, these factors could not be included in the analyses.

It has been suggested that adjuvant therapy could improve oncological outcome in patients with lymph node metastases. However, a meta-analysis[40] found that patients with rectal cancer did not benefit from adjuvant chemotherapy with regard to DFS (HR 0.91, 95 per cent c.i. 0.77 to 1.07; P=0.230) and distant recurrence (HR 0.94, 0.78 to 1.14; P=0.523) compared with observation. In the present study, cN category lacked predictive value for survival outcomes. This was probably related to the low accuracy of clinical nodal staging, which was mainly performed with endorectal ultrasonography and CT. Currently, MRI is the recommended modality for assessment of node status; however, T2-weighted MRI also only yields a moderate sensitivity and specificity of 77 and 60 per cent respectively. The per-lesion sensitivity for nodal staging after chemoradiotherapy is 91 per cent, indicating a low rate of false-negative findings when staging individual mesorectal nodes.[41] Lahaye and colleagues[42] reported sensitivities of up to 85 per cent for nodal staging with T2- weighted MRI after chemoradiotherapy based on size criteria, further confirming the low risk of missing lymph node metastases. Nevertheless, given the 17 per cent prevalence of lymph node metastases in ypT2 disease, physicians should remain alert to the possible presence of lymph node metastases in patients with substantial downstaging of the primary rectal cancer.

This study has several limitations. Data were retrieved from a subset of individual studies with a heterogeneous patient population and differences between studies. Some of the studies were retrospective. However, a random-effects model was used to take heterogeneity into account when pooling the proportions of lymph node metastases by ypT category, and Cox proportional hazards analyses with stratification by data set were used to evaluate long-term outcome. Because of missing data, not all patients could be included in all analyses. Additionally, some baseline and histopathological details were lacking, such as the presence of tumour deposits, extramural vascular invasion, completeness of resection, size and number of harvested and involved nodes, and size and exact location of residual tumour in the bowel wall; this information could be of help in interpreting the data.[38, 39, 43] Moreover, clinical staging was probably suboptimal (specifically for nodal status) as MRI was not used in most studies, which may have influenced the outcomes. Finally, this pooled analysis was

based on historical studies published between 2002 and 2008. However, this provided a unique opportunity to evaluate lymph node metastases in patients with rectal cancer who receive chemoradiotherapy and all undergo surgery, in contrast to current cohorts in which organ reservation is increasingly being offered.

ACKNOWLEDGEMENTS

J. Garcia-Aguilar's work on this study was funded in part by grant P30 CA008748 from the National Cancer Institute.

SUPPLEMENTARY MATERIAL

	Valentini[22]	Rödel[28]	Kuo[26]	García-Aguilar[19]	Glynne-Jones[20]	Suárez[27]	Calvo[21]	Pucciarelli[23]	Biondo[25]	Theodoropoulos[24]
N	474	348	242	154	147	119	117	106	103	88
Mean age (range)	61 (22-83)	67 (38-85)	58 (25-87)	60 (26-83)	63 (21-83)	68 (41-87)	65 (26-85)	60 (31-79)	59 (20-77)	63 (29-83)
Gender (male)	61% (293/474)	71% (246/348)	60% (144/242)	56% (86/154)	71% (103/146)	70% (83/119)	67% (78/117)	64% (68/106)	73% (75/103)	72% (63/88)
Missing	0%	0%	0%	0%	1% (1/147)	0%	0%	0%	0%	0%
Median FU	53 (1-219)	91 (3-156)	33 (1-121)	32 (1-111)	36 (0-140)	42 (3-86)	71 (0-143)	61 (9-141)	41 (0-94)	33 (2-95)
months (range)										
Clinical T-stage										
T1	0% (2/447)	0%	0%	1% (2/154)	0%	0%	0%	0%	0%	0%
T2	4% (16/447)	5% (14/278)	6% (15/242)	5% (8/154)	4% (6/145)	7% (8/118)	3% (3/116)	11% (11/104)	2% (2/103)	0%
T3	83% (369/447)	87% (243/278)	85% (206/242)	83% (128/154)	48% (69/145)	88% (104/118)	87% (101/116)	70% (73/104)	92% (95/103)	94% (83/88)
T4	13% (60/447)	8% (21/278)	9% (21/242)	11% (16/154)	48% (70/145)	5% (6/118)	10% (12/116)	19% (20/104)	6% (6/103)	6% (5/88)
Missing	6% (27/474)	20% (70/348)	0%	0%	1% (2/147)	1% (1/119)	1% (1/117)	2% (2/106)	0%	0%
Clinical N-stage										
N0	19% (94/448)	82% (268/327)	14% (30/209)	29% (44/154)	47% (68/144)	58% (68/118)	52% (59/113)	35% (36/103)	19% (19/98)	52% (46/88)
N+	81% (364/448)	18% (59/327)	86% (179/209)	71% (110/154)	53% (76/144)	42% (50/118)	48% (54/113)	65% (67/103)	81% (79/98)	48% (42/88)
Missing	5% (26/474)	6% (21/348)	14% (33/242)	0%	2% (3/147)	1% (1/119)	3% (4/117)	3% (3/106)	5% (5/103)	0%
Surgery										
LAR	76% (358/474)	68% (237/348)	82% (199/242)	60% (92/154)	37% (55/147)	74% (88/119)	59% (69/117)	87% (92/106)	60% (62/103)	64% (56/88)
APR	18% (84/474)	25% (88/348)	14% (34/242)	38% (58/154)	61% (89/147)	23% (28/119)	41% (48/117)	12% (13/106)	40% (41/103)	36% (32/88)
Other	7% (32/474)	7% (23/348)	4% (9/242)	2% (4/154)	2% (3/147)	3% (3/119)	0%	1% (1/106)	0%	0%
Missing	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Response										
pCR	19% (92/473)	10% (34/344)	15% (36/236)	14% (21/151)	18% (26/143)	14% (17/119)	8% (10/117)	18% (19/106)	15% (16/103)	19% (17/88)
ypT1-2	35% (162/473)	38% (131/344)	30% (70/236)	35% (53/151)	21% (30/143)	30% (35/119)	44% (51/117)	50% (53/106)	37% (38/103)	22% (19/88)
ypT3-4	46% (219/473)	52% (179/344)	55% (130/236)	51% (77/151)	61% (87/143)	56% (67/119)	48% (56/117)	32% (34/106)	48% (49/103)	59% (52/88)
Missing	0% (1/474)	1% (4/348)	2% (6/242)	2% (3/154)	3% (4/147)	0%	0%	0%	0%	0%
pN status										
ypN0	73% (348/474)	70% (244/348)	67% (163/242)	69% (107/154)	75% (110/147)	76% (91/119)	80% (94/117)	81% (86/106)	75% (77/103)	67% (59/88)
ypN+	27% (126/474)	30% (104/348)	33% (79/242)	31% (47/154)	25% (37/147)	24% (28/119)	20% (23/117)	19% (20/106)	25% (26/103)	33% (29/88)
Missing	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Adjuvant chemotherapy	21% (98/474)	85% (295/348)	60% (145/242)	21% (32/154)	20% (29/147)	82% (97/119)	62% (73/117)	57% (60/106)	100% (103/103)	56% (49/88)

Table S1 Baseline characteristics of the patients per included dataset. FU=follow-up, LAR=low anterior resection; APR=abdominoperineal resection; pCR=pathologic complete response

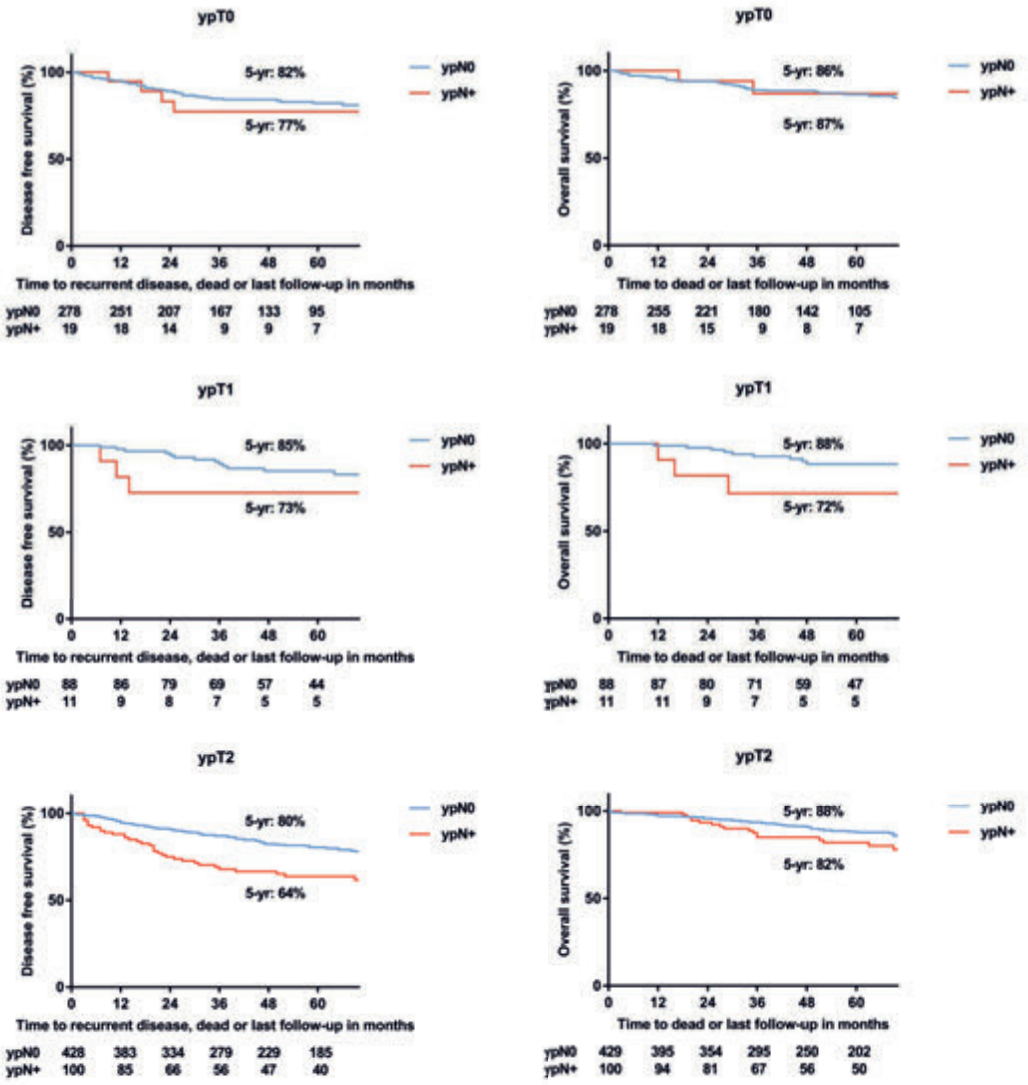


Figure S1. Survival curves for DFS and OS by ypN status separated by ypT category. Percentages are the 5-year DFS and OS per group.

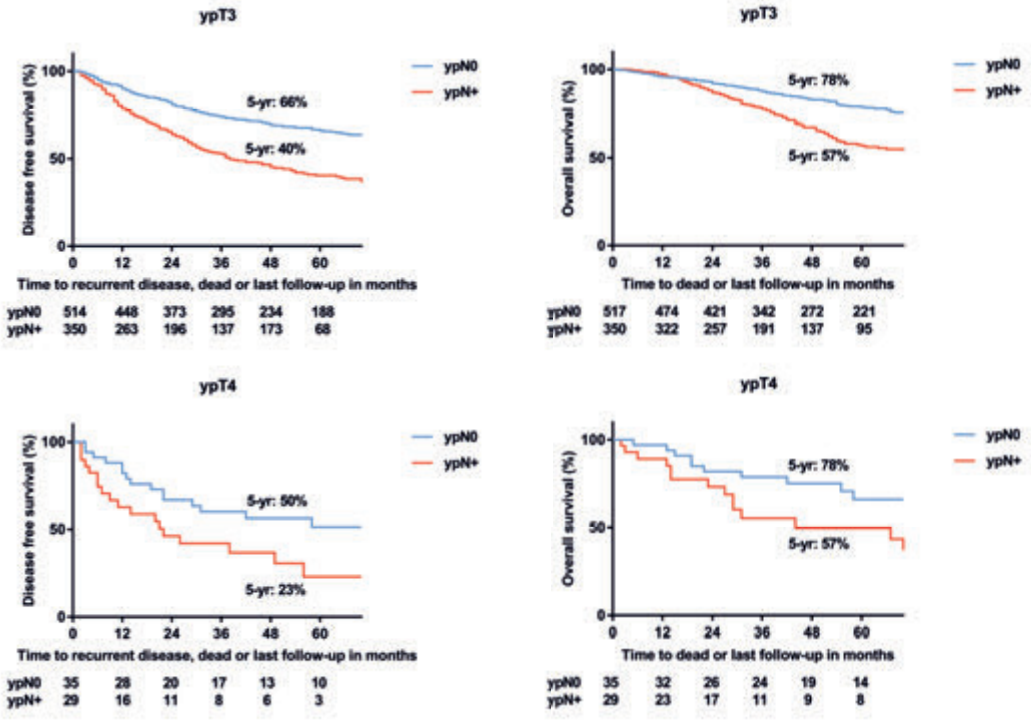


Figure S1. Continued.

REFERENCES

1. Heald RJ, Ryall RD. Recurrence and survival after total mesorectal excision for rectal cancer. *Lancet*. Jun 28 1986;1(8496):1479-82.
2. Sauer R, Becker H, Hohenberger W, et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. *N Engl J Med*. Oct 21 2004;351(17):1731-40.
3. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
4. Dossa F CTR, Acuna S.A, Baxter N.N. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete respons following neoadjuvant chemoradiation: a systematic review and meta-analysis. *Lancet Gastroenterology heptology*. 2017;(2):501-13.
5. Dattani M, Heald RJ, Goussous G, et al. Oncological and Survival Outcomes in Watch and Wait Patients With a Clinical Complete Response After Neoadjuvant Chemoradiotherapy for Rectal Cancer: A Systematic Review and Pooled Analysis. *Ann Surg*. 2018 Dec;268(6):955-967.
6. Rullier E, Vendrely V, Asselineau J, et al. Organ preservation with chemoradiotherapy plus local excision for rectal cancer: 5-year results of the GRECCAR 2 randomised trial. *Lancet Gastroenterol Hepatol*. May 2020;5(5):465-474.
7. Borschitz T, Wachtlin D, Mohler M, Schmidberger H, Junginger T. Neoadjuvant chemoradiation and local excision for T2-3 rectal cancer. *Ann Surg Oncol*. Mar 2008;15(3):712-20.
8. Ricciardi R, Madoff RD, Rothenberger DA, Baxter NN. Population-based analyses of lymph node metastases in colorectal cancer. *Clin Gastroenterol Hepatol*. Dec 2006;4(12):1522-7.
9. Blumberg D, Paty PB, Guillem JG, et al. All patients with small intramural rectal cancers are at risk for lymph node metastasis. *Dis Colon Rectum*. Jul 1999;42(7):881-5.
10. Fang WL, Chang SC, Lin JK, et al. Metastatic potential in T1 and T2 colorectal cancer. *Hepatogastroenterology*. Nov-Dec 2005;52(66):1688-91.
11. Stornes T, Wibe A, Nesbakken A, Myklebust T, Endreseth BH. National Early Rectal Cancer Treatment Revisited. *Dis Colon Rectum*. Jul 2016;59(7):623-9.
12. Rasheed S, Bowley DM, Aziz O, et al. Can depth of tumour invasion predict lymph node positivity in patients undergoing resection for early rectal cancer? A comparative study between T1 and T2 cancers. *Colorectal Dis*. Mar 2008;10(3):231-8.
13. Sitzler PJ, Seow-Choen F, Ho YH, Leong AP. Lymph node involvement and tumor depth in rectal cancers: an analysis of 805 patients. *Dis Colon Rectum*. Dec 1997;40(12):1472-6.
14. Yamamoto S, Watanabe M, Hasegawa H, et al. The risk of lymph node metastasis in T1 colorectal carcinoma. *Hepatogastroenterology*. Jul-Aug 2004;51(58):998-1000.
15. Nagtegaal ID, Marijnen CAM. The future of TNM staging in rectal cancer: The era of neoadjuvant therapy. *Curr Colorectal Cancer Rep*. 2008/07/01 2008;4(3):147-154.
16. Mignanelli ED, de Campos-Lobato LF, Stocchi L, Lavery IC, Dietz DW. Downstaging after chemoradiotherapy for locally advanced rectal cancer: is there more (tumor) than meets the eye? *Dis Colon Rectum*. Mar 2010;53(3):251-6.

17. Rullier E, Rouanet P, Tuech J-J, et al. Organ preservation for rectal cancer (GRECCAR 2): a prospective, randomised, open-label, multicentre, phase 3 trial. *The Lancet*. 2017;390(10093):469-479.
18. Maas M, Nelemans PJ, Valentini V, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol*. Sep 2010;11(9):835-44.
19. García-Aguilar J, Hernandez de Anda E, Sirivongs P, Lee SH, Madoff RD, Rothenberger DA. A pathologic complete response to preoperative chemoradiation is associated with lower local recurrence and improved survival in rectal cancer patients treated by mesorectal excision. *Dis Colon Rectum*. Mar 2003;46(3):298-304.
20. Hughes R, Glynne-Jones R, Grainger J, et al. Can pathological complete response in the primary tumour following pre-operative pelvic chemoradiotherapy for T3-T4 rectal cancer predict for sterilisation of pelvic lymph nodes, a low risk of local recurrence and the appropriateness of local excision? *Int J Colorectal Dis*. Jan 2006;21(1):11-7.
21. Díaz-González JA, Calvo FA, Cortés J, et al. Prognostic factors for disease-free survival in patients with T3-4 or N+ rectal cancer treated with preoperative chemoradiation therapy, surgery, and intraoperative irradiation. *Int J Radiat Oncol Biol Phys*. Mar 15 2006;64(4):1122-8.
22. Valentini V, Coco C, Picciocchi A, et al. Does downstaging predict improved outcome after preoperative chemoradiation for extraperitoneal locally advanced rectal cancer? A long-term analysis of 165 patients. *Int J Radiat Oncol Biol Phys*. Jul 1 2002;53(3):664-74.
23. Pucciarelli S, Toppan P, Friso ML, et al. Complete pathologic response following preoperative chemoradiation therapy for middle to lower rectal cancer is not a prognostic factor for a better outcome. *Dis Colon Rectum*. Nov 2004;47(11):1798-807.
24. Theodoropoulos G, Wise WE, Padmanabhan A, et al. T-level downstaging and complete pathologic response after preoperative chemoradiation for advanced rectal cancer result in decreased recurrence and improved disease-free survival. *Dis Colon Rectum*. Jul 2002;45(7):895-903.
25. Biondo S, Navarro M, Marti-Rague J, et al. Response to neoadjuvant therapy for rectal cancer: influence on long-term results. *Colorectal Dis*. Sep 2005;7(5):472-9.
26. Kuo LJ, Liu MC, Jian JJ, et al. Is final TNM staging a predictor for survival in locally advanced rectal cancer after preoperative chemoradiation therapy? *Ann Surg Oncol*. Oct 2007;14(10):2766-72.
27. Suárez J, Vera R, Balén E, et al. Pathologic response assessed by Mandard grade is a better prognostic factor than down staging for disease-free survival after preoperative radiochemotherapy for advanced rectal cancer. *Colorectal Dis*. Jul 2008;10(6):563-8.
28. Rödel C, Martus P, Papadoupolos T, et al. Prognostic significance of tumor regression after preoperative chemoradiotherapy for rectal cancer. *J Clin Oncol*. Dec 1 2005;23(34):8688-96.
29. Nyaga VN, Arbyn M, Aerts M. Metaprop: a Stata command to perform meta-analysis of binomial data. *Arch Public Health*. 2014;72(1):39.
30. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled clinical trials*. Sep 1986;7(3):177-88.
31. Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions. Cochrane. 2020, 20th of July Accessed 20th of July 2020, version 6.0. www.training.cochrane.org/handbook

32. Erkan A, Mendez A, Trepanier M, et al. Impact of residual nodal involvement after complete tumor response in patients undergoing neoadjuvant (chemo)radiotherapy for rectal cancer. *Surgery*. Oct 2019;166(4):648-654.
33. Duchalais E, Glyn Mullaney T, Spears GM, et al. Prognostic value of pathological node status after neoadjuvant radiotherapy for rectal cancer. *Br J Surg*. Oct 2018;105(11):1501-1509.
34. Calvo FA, Morillo V, Santos M, et al. Interval between neoadjuvant treatment and definitive surgery in locally advanced rectal cancer: impact on response and oncologic outcomes. *J Cancer Res Clin Oncol*. Oct 2014;140(10):1651-60.
35. Ryan É J, O'Sullivan DP, Kelly ME, et al. Meta-analysis of the effect of extending the interval after long-course chemoradiotherapy before surgery in locally advanced rectal cancer. *Br J Surg*. Sep 2019;106(10):1298-1310.
36. Akgun E, Caliskan C, Bozbiyik O, et al. Randomized clinical trial of short or long interval between neoadjuvant chemoradiotherapy and surgery for rectal cancer. *Br J Surg*. Oct 2018;105(11):1417-1425.
37. Detering R, van Oostendorp SE, Meyer VM, et al. MRI cT1-2 rectal cancer staging accuracy: a population-based study. *Br J Surg*. Sep 2020;107(10):1372-1382.
38. Smith FM, Chang KH, Sheahan K, Hyland J, O'Connell PR, Winter DC. The surgical significance of residual mucosal abnormalities in rectal cancer following neoadjuvant chemoradiotherapy. *Br J Surg*. Jul 2012;99(7):993-1001.
39. Bosch SL, Vermeer TA, West NP, et al. Clinicopathological characteristics predict lymph node metastases in ypT0-2 rectal cancer after chemoradiotherapy. *Histopathology*. Nov 2016;69(5):839-848.
40. Breugom AJ, Swets M, Bosset JF, et al. Adjuvant chemotherapy after preoperative (chemo) radiotherapy and surgery for patients with rectal cancer: a systematic review and meta-analysis of individual patient data. *Lancet Oncol*. Feb 2015;16(2):200-7.
41. van der Paardt MP, Zagers MB, Beets-Tan RG, Stoker J, Bipat S. Patients who undergo preoperative chemoradiotherapy for locally advanced rectal cancer restaged by using diagnostic MR imaging: a systematic review and meta-analysis. *Radiology*. Oct 2013;269(1):101-12.
42. Lahaye MJ, Beets GL, Engelen SM, et al. Locally advanced rectal cancer: MR imaging for restaging after neoadjuvant radiation therapy with concomitant chemotherapy. Part II. What are the criteria to predict involved lymph nodes? *Radiology*. Jul 2009;252(1):81-91.
43. Nougaret S, Castan F, de Forges H, et al. Early MRI predictors of disease-free survival in locally advanced rectal cancer from the GRECCAR 4 trial. *Br J Surg*. Oct 2019;106(11):1530-1541.

CHAPTER 4

The evaluation of follow-up strategies of watch-and-wait patients with a complete response after neoadjuvant therapy in rectal cancer

Hester E. Haak, Jan Žmuc, Doenja M. J. Lambregts, Regina G. H. Beets-Tan, Jarno Melenhorst, Geerard L. Beets, Monique Maas, on behalf of the Dutch Watch-and-Wait Consortium

Colorectal Dis. 2021 Jul;23(7):1785-1792

ABSTRACT

Aim

Many of the current follow-up schedules in a watch-and-wait approach include very frequent MRI and endoscopy examinations to ensure early detection of local regrowth (LR). The aim of this study was to analyze the occurrence and detection of LR in a watch-and-wait cohort and to suggest a more efficient follow-up schedule.

Method

Rectal cancer patients with a clinical complete response after neoadjuvant therapy were prospectively and retrospectively included in a multicenter watch-and-wait registry between 2004 and 2018, with the current follow-up schedule with 3-monthly endoscopy and MRI in the first year and 6-monthly thereafter. A theoretical comparison was constructed for the detection of LR in the current follow-up schedule against four other hypothetical schedules.

Results

In all, 50/304 (16%) of patients developed a LR. The majority (98%) were detected at ≤2 years, located in the lumen (94%) and were visible on endoscopy (88%). The theoretical comparison of the different hypothetical schedules suggests that the most optimal follow-up schedule should focus on the first 2 years with 3-monthly endoscopy and 3-6 monthly MRI. Longer intervals in the first 2 years will cause delays in diagnosis of LR ranging from 0-5 months. After 2 years, increasing the interval from 6 to 12 months did not cause important delays.

Conclusion

The most optimal follow-up schedule for a watch-and-wait policy in patients with a clinical complete response after chemoradiation for rectal cancer should include frequent endoscopy and to a lesser degree MRI in the first 2 years. Longer intervals, up to 12 months, can be considered after 2 years.

INTRODUCTION

During the last decade, the watch-and-wait (W&W) approach has been accepted as an alternative treatment in rectal cancer patients with a clinical complete response (cCR) after neoadjuvant therapy.[1-3] Adequate follow-up in W&W patients is essential for early detection and treatment of local regrowths (LRs), in order to achieve similar long-term outcomes compared to patients who undergo a standard rectal resection. It has been widely accepted that a three-modality approach has the highest accuracy to detect complete responders with frequent digital rectal examination (DRE), endoscopy and MRI with diffusion-weighted-imaging (DWI)[1]. Most centers agree on a more frequent surveillance during the first 2 years, but there is a marked difference in the schedules regarding frequency and use of MRI and endoscopy. [1, 4] Intensive follow-up visits and examinations can be a burden for patients, especially the frail and elderly. In addition, there is little information on the efficiency of frequent follow-up examinations, and on the value of MRI and endoscopy in detecting LR. In order to improve W&W follow-up, there is a need to balance between optimal LR detection, burden and efficiency. The current intensive follow-up protocol in the Dutch W&W network was based more on safety concerns than on evidence. The aim of this study is to analyze the occurrence and detection of LR in a W&W cohort and to suggest a more efficient follow-up schedule.

METHODS

Details of the W&W programme

Patients diagnosed with rectal cancer who had a cCR after neoadjuvant therapy who were offered a W&W programme between 2004 and 2017 were prospectively included in a local study from the Maastricht University Medical Center, approved by the local institutional review board and registered in clinicaltrials.gov since 2009 (NCT00939666 and NTC02278653), and provided informed consent. W&W patients from 2017 to 2018 were retrospectively included in a quality-controlled national registration of W&W patients, for which informed consent was waived by the local institutional review board. Patients were included in a W&W programme if they had a biopsy proven rectal adenocarcinoma without distant metastasis at baseline and received neoadjuvant treatment with long course chemoradiation consisting of 28x1.8 Gy with 2x825 mg/m^3 capecitabine or short course radiotherapy with 5 x 5 Gy followed by a waiting interval. Patients underwent restaging approximately 8-12 weeks after completion of (chemo)radiation by digital rectal examination (DRE), endoscopy and MRI including diffusion weighted imaging (MRI-DWI). Those who were identified during restaging with a cCR or patients with a near complete response (nCR) were included in W&W. A cCR was defined as (1) no residual tumour felt on DRE, (2) white scar and/or telangiectasia of the mucosa on endoscopy and (3) low signal intensity at the original tumour site on T2 weighted MRI with absence of diffusion restriction on MRI-DWI and absence of residual malignant nodes. [5, 6] A nCR was defined as (1) minor soft mucosal

abnormality or irregularity felt on DRE, (2) superficial ulceration and/or mild persisting erythema of the scar and (3) intermediate or low residual signal on T2-weighted MRI and/or small foci of diffusion restriction on MRI-DWI. [5, 6] All patients included for W&W were informed of the experimental nature of the study and were aware that the W&W approach was an alternative treatment and deviated from current guidelines. The current follow-up schedule in the Dutch hospital network consists of 3-monthly endoscopy and MRI in the first year and 6-monthly thereafter. [7] Standard follow-up methods for distant metastasis (DM) consisted of CT imaging of the chest and liver and CEA blood levels for 5 years, according to national guidelines. [8]

Study cohort for the analysis of detection of regrowths

First, we analysed the timing and modality of regrowths. Patients who were included in the W&W programme and who developed a LR during follow-up were eligible for the analysis of detection of regrowths. In order to provide a strictly selected study cohort, W&W patients who developed a typical cCR on MRI and endoscopy at first or second restaging (after another 6- to 12-week interval) were selected and W&W patients with a persisting nCR at second restaging or patients with local excision (TEM) prior to inclusion for W&W were excluded. Although it was intended that patients followed the advised current follow-up schedule (3-monthly endoscopy and MRI in the first year and 6-monthly thereafter), in reality, some patients had fewer examinations while others had more frequent examinations because of patient preference, logistical planning issues or findings on endoscopy and/or MRI that warranted earlier follow-up. These variations could be used to evaluate the delay of LR detection in the current follow-up schedule and allowed to also study more intensive hypothetical schedules. The detection of LR with the actual follow-up schedule in the study cohort was compared with the estimated timing of regrowth detection if the current follow-up schedule would have been followed and in four additional hypothetical follow-up schedules. At the start of the study, before any analysis was performed, the study group agreed on the four hypothetical schedules, based on literature and own experience. Because many studies have shown a low incidence of LRs after two years, all four hypothetical follow-up schedules consisted of less frequent examinations after two years [1-3]. For the first two years two schedules tested more frequent, and two schedules less frequent examinations. The hypothetical schedules were as follows:

1. Schedule 1: 3-monthly endoscopy and MRI in the first year, 3-monthly endoscopy and 6-montly MRI in the second year and yearly endoscopy and MRI thereafter.
2. Schedule 2*:* 3-monthly endoscopy and MRI in the first year and 4-monthly in the second year and yearly endoscopy and MRI thereafter.
3. Schedule 3: 4-monthly endoscopy and MRI in the first year and 6-monthly in the second year and yearly endoscopy and MRI thereafter.
4. Schedule 4: 6-monthly endoscopy and MRI during the first 2 years and yearly endoscopy and MRI thereafter.

Comparing detection of regrowths in different follow-up schedules

To identify the most optimal schedule, the actual LR detection, defined as the LR detection according to the actually performed evaluations in the study cohort was compared with the estimated LR detection in the current and hypothetical schedules. Delay in LR detection was calculated as the difference between the actual LR detection and LR detection in the current and hypothetical schedules. For the analyses, there were two assumptions. The first was that when the examination with which the LR was actually detected in the series was left out in a theoretical schedule, the regrowth would be detected at the next scheduled examination. The second assumption was that in a theoretical schedule a regrowth cannot be detected earlier than when it was actually detected in the series.

Statistical analyses

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS version 25.0). Baseline data were collected for all patients and included age, sex, baseline clinical staging, neoadjuvant therapy, type of surgical procedure, adjuvant chemotherapy and median follow-up time. Quantitative data were expressed as median with a range of minimum and maximum values. Categorical data were reported as the number of patients with percentages. LR was defined as tumour regrowth in the lumen or in mesorectal lymph nodes. Duration of follow-up and interval to event were calculated from the date of restaging MRI to the event of interest or last follow-up date that was used as a date of censoring.

RESULTS

Demographics

Figure 1 shows a flowchart with an overview of included and excluded patients. Fifty (16%) of 304 patients developed a LR during follow-up with a 2-year LR rate of 17%. 23 (46%) of 50 LR patients had a cCR during restaging and 27 (54%) had a nCR during restaging but achieved a cCR at second restaging. Median age of LR patients was 64 years (range 43-85). Of the 50 LR patients, 42 (84%) had a distal tumour (≤ 5 cm of the anorectal junction) and 8 (16%) had a mid-rectum tumour (5.1 – 10 cm of the anorectal junction). Median follow-up time was 30 months (9-115) and median time from end of radiotherapy to date of restaging MRI was 9 weeks (5-18). A more detailed overview of baseline characteristics of W&W patients and those who developed a LR and who were eligible for the analysis are shown in Table 1.

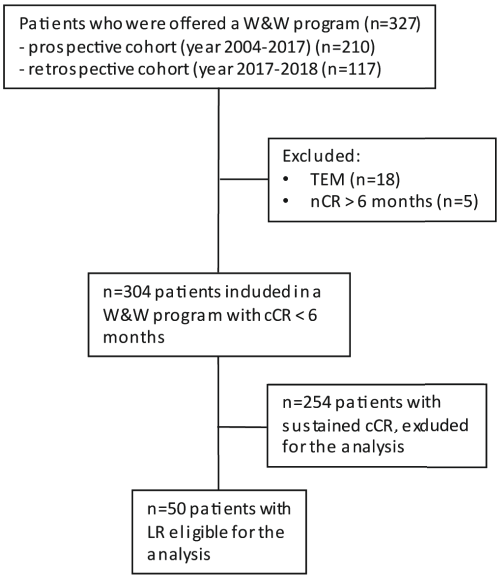


Figure 1 Flowchart with an overview of included and excluded patients. cCR, clinical complete response; LR, local regrowth; nCR, near complete response; TEM, transanal endoscopic microsurgery; W&W, watch-and-wait

	W&W patients (n=304)	Patients with LR (n=50) eligible for analysis
Median age (years)	66 (33-87)	64 (43-85)
Sex (male)	67% (204/304)	72% (36/50)
Clinical T stage		
T1	1% (2/304)	0% (0/0)
T2	21% (66/304)	10% (5/50)
T3	69% (209/304)	74% (37/50)
T4	9% (27/304)	16% (8/50)
Clinical N stage (N+)	73% (222/304)	68% (34/50)
Distance anal verge (cm)		
<5	78% (237/304)	84% (42/50)
>5	22% (67/304)	16% (8/50)
Neoadjuvant therapy		
CRT	94% (285/304)	96% (48/50)
5x5Gy with a long waiting interval	5% (16/304)	4% (2/50)
Other	1% (3/304)	NA
Adjuvant chemotherapy	16% (47/304)	10% (5/50)

Table 1 Baseline characteristics of W&W patients, and those with a local regrowth included for analyses to evaluate different follow-up schedules. Data are median (range) or %(n/N). CRT, chemoradiotherapy; LR, local regrowth

Patients with a local regrowth

The majority of LRs were diagnosed within 2 years (n=49, 98%). The only patient with a regrowth later than 2 years had a nodal regrowth along the superior rectal vessels at the level of L5 diagnosed after 3 years and 8 months. In retrospect this was missed at MRI and the node was already visible 21 months earlier on MRI after 23 months of follow-up (i.e. this was in fact also a recurrence within 2 years). LRs were located luminal-only in 42 (84%) patients, both luminal and nodal in five (10%), and in regional lymph nodes only in three (6%) (Figure 2, 3 and 4). The majority were detected on both endoscopy and MRI (n=32, 64%), in 12 (24%) only on endoscopy and in six (12%) LR was only detected on MRI.

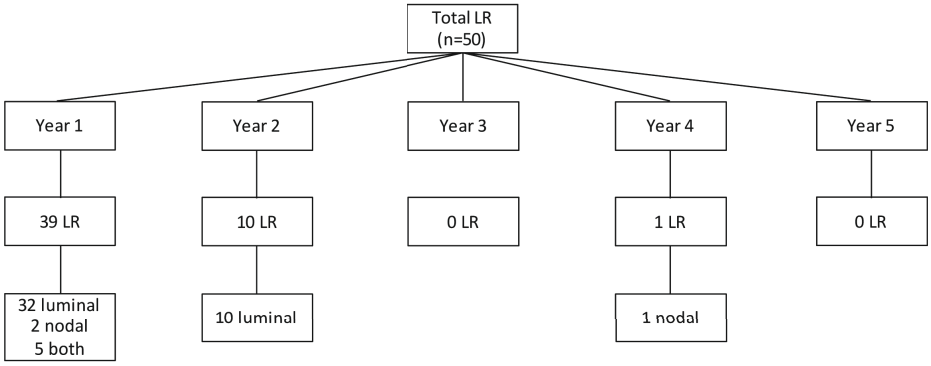


Figure 2 Flowchart of local regrowths. LR, local regrowth

Comparing detection of regrowths in different FU schedules

The current follow-up schedule used in the Dutch hospital network consists of 24 examinations (12 endoscopy with DRE and 12 MRI-DWI) in 5 years after the inclusion in the W&W programme. Because some patients had more follow-up examinations than required in the standard protocol because of patient preference or logistical planning issue or findings on endoscopy and/or MRI that warranted earlier follow-up, some LRs were actually detected ahead of the standard assessment date. Supplementary table 1 provides a detailed overview of all patients with a LR and the theoretical difference in detection time-point according to the current and hypothetical schedules. The overall median delay in LR detection was 0 (range 0-5) months for the current schedule and 0 (range 0-4), 0 (range 0-4), 0 (range 0-4) and 2 (range 0-5) months for hypothetical schedule 1, 2, 3 and 4, respectively.

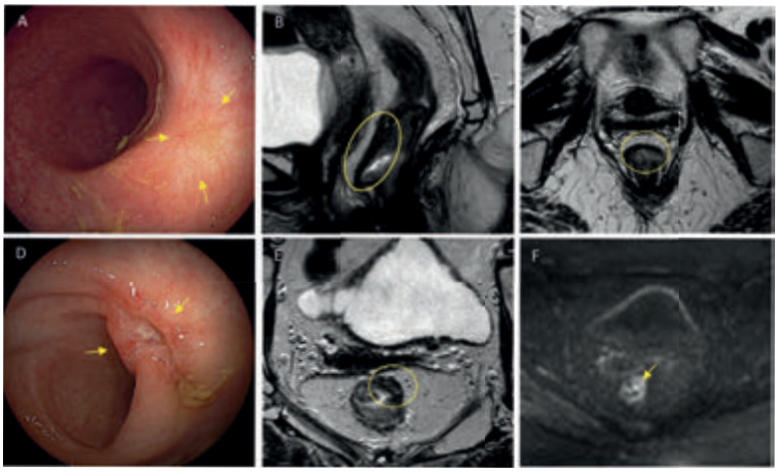


Figure 3 Example of patient with rectal cancer with a clinical complete response after neoadjuvant treatment. (A) White scar tissue and telangiectasia (yellow arrows) on endoscopy and (B) corresponding fibrotic wall on sagittal and (C) transversal T2-weighted MR images (indicated in yellow). Six months later, (D) there is an ulcer with elevated edges on endoscopy (yellow arrows) and (E) tumour mass is visible on transversal T2-weighted MR images within the fibrotic tumour bed with (F) diffusion restriction on diffusion-weighted imaging, suspicious for a local regrowth.

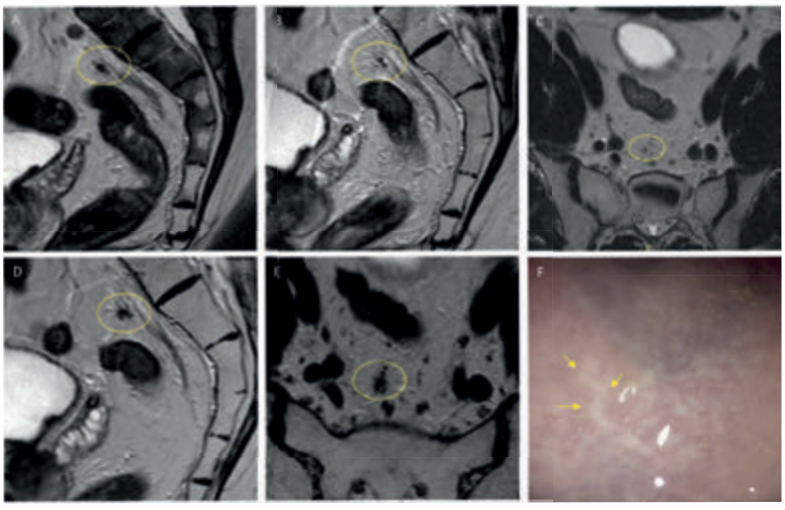


Figure 4 Example of a patient with rectal cancer with a malignant lymph node on (A) sagittal T2-weighted MR images before chemoradiation treatment. After chemoradiation treatment (B), (C) the lymph node decreased in size and was considered as no longer suspect. 12 months later (D), (E) the lymph node has grown, suggestive of nodal regrowth, while maintaining a luminal complete response on endoscopy (yellow arrows) (F)

Figure 5 provides an overview of all patients with at least 3 months of delay in LR detection with both the current follow-up schedule and hypothetical schedules. In addition to the current follow-up schedule (24 examinations), the four hypothetical schedules consisted of 20, 20, 16 and 14 examinations for schedule 1, 2, 3 and 4, respectively. With the current follow-up schedule, four patients with at least 3 months of delay in detection of LRs occurred in the first 2 years of follow-up. Hypothetical schedule 1 would have zero delays of at least 3 months during the first 2 years of follow-up, schedule 2 would have two delays of at least 3 months, schedule 3 would have 11 delays of at least 3 months and schedule 4 would have 14 delays of at least 3 months. In both the current schedule and the hypothetical schedules, one delay of at least 3 months in LR detection occurred after 2 years of follow-up in the patient described above with a nodal regrowth that was detected at 3 years and 8 months, but that was in retrospect visible at 21 months.

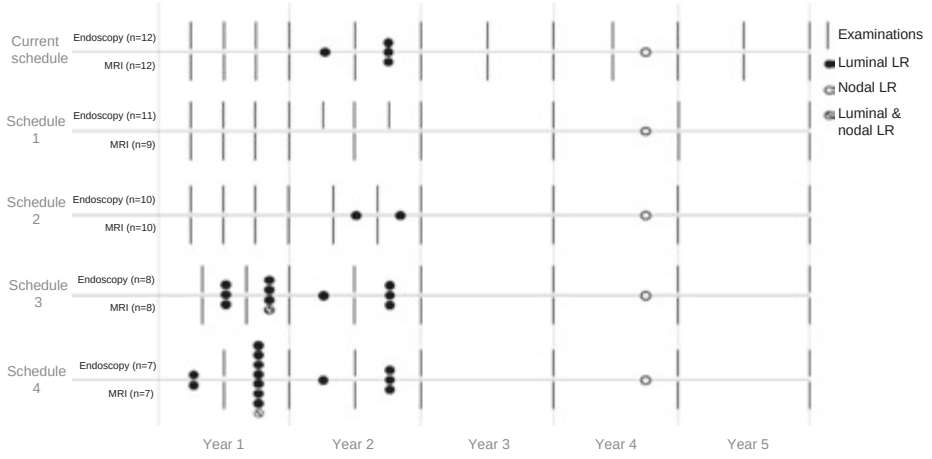


Figure 5 Overview of all patients with at least 3 months of delay in LR detection with both the current follow-up schedule and hypothetical schedules. LR, local regrowths; n, number of examinations

DISCUSSION

The majority of the LRs in a W&W approach for a complete response after neoadjuvant therapy of rectal cancer were detected within 2 years (98%), located in the bowel wall (94%) and were visible on endoscopy (88%). The optimal follow-up schedule focuses on the first 2 years, with an intensive follow-up including 3-monthly combined endoscopy and MRI assessment in the first year. In the second year the MRI can be performed at a 6-monthly interval combined with endoscopy every 3 months. This schedule minimizes the delay in detection of regrowths based on the available outcome data in the current series. Moreover, this schedule de-intensified the current follow-up schedule from 24 examinations to 20 examinations. De-intensifying the follow-up examinations in the first 2 years (schedules 3

and 4) resulted in more delays. Because very few regrowths became evident after two years, the follow-up interval can be de-intensified in years 3-5, that is, to 12-monthly follow-up, with no extra delay in detection.

Other studies also reported that most regrowths are luminal [1, 3], and a number of W&W centers mainly rely on frequent endoscopies during the first 2 years. [9-11] It is known that clinical assessment with DRE and endoscopy is the single most accurate modality for identification of complete responders.[4] The most commonly used endoscopic technique is standard high-resolution endoscopy with white light. There are several new endoscopic techniques using advanced imaging such as narrow band imaging and chromoendoscopy which may improve the diagnostic accuracy of endoscopy in the future.[12, 13] However, more studies need to confirm its added value before these techniques will be implemented in a W&W follow-up. The policy in most centers is to rely on serial endoscopic assessments, and perform targeted biopsies of any changes in the scar. When adenocarcinoma is found the interpretation is easy, but there is always the risk of a false negative biopsy through sampling error, and adenomatous changes and high-grade dysplasia can be difficult to interpret. [5, 14, 15]

Our finding of the vast majority of regrowths occurring in the first 2 years of follow-up has also been noted by others [1-3], and de-intensifying the follow-up interval after 2 years has been recommended before. [1] Some groups even minimize the follow-up after 2 years to standard surveillance with regular CT scans and CEA measurements and omission of specific W&W follow-up. [16, 17] Moreover, recent updated Dutch guidelines even recommend to reduce the standard surveillance to regular CEA measurements and only perform CT scans by indication.[18] Considering the low risk of regrowths after 2 years and only one (discovered late) false negative finding in the current study, some groups may even opt to further reduce the number of examinations after 2 years which will increase the cost-effectiveness. Although it is clear that the efficiency of regular follow-up with MRI and endoscopy is lower after two years, there is a small number of patients who could benefit from early detection of a late regrowth. While we propose a yearly follow-up after two years of follow-up, some less experienced centres may feel more comfortable with a more gradual decrease as more assessments can compensate for missed detections, for example by maintaining a 6-monthly interval in year 3 and going to a 12-monthly follow-up in year 4 and 5. The single patient in our study with a late regrowth after 2 years had a high nodal deposit while the luminal tumour was still in complete remission. In retrospect the growing node was already visible on MRI scans more than a year earlier (after 23 months of follow-up) and the nodal regrowth was visible at CT as well, highlighting the importance of both the technical quality of the MRI as well as attentive reading by radiologists. The field of view of the MRI (both sagittal and axial) has to be wide enough to encompass the lateral nodal area as well as the proximal nodal area at the level of the promontory. Even though 18F-fluorodeoxyglucose PET (PET) might help in detection of malignant nodes, its use as part of the standard routine is unlikely, given the costs and availability. It can be an adjunct when in doubt about nodes or other potential tumour metastases, e.g. in case of an increased CEA.[19-21] In addition, as late nodal

regrowths are rare, standard follow-up with CT for distant metastasis also aids in detecting these regrowths, which makes the risk of missed regrowths due to de-intensification of follow-up schedules small.

This study has several limitations. First, the number of LRs was relatively small. Second, although the majority of patients were prospectively registered, some of the details of the endoscopy and MRI reports, such as modality of detection of a LR and the interpretation by the clinician had to be identified and interpreted retrospectively, which could have caused minor issues in determining the exact timing and modality of diagnosis of the regrowth. Third, it has to be noted that duration of follow-up and interval to event was calculated from date of restaging MRI. This has to be taken into account when comparing the results to studies with different starting points, such as the start or end of radiotherapy. Fourth, for various reasons, for example findings that needed short follow-up (e.g. change on MRI), some patients had more frequent examinations. Patients who correctly followed the current protocol could not be taken into account to evaluate this more intensive FU schedule, which leads to measurement bias. Last, patients were included in centers with experience in W&W and caution is required when extrapolating results to those from less experienced centers.

This study provides an overview of LRs during W&W that can be used to adapt the current strict follow-up protocol for W&W. The results support an intensive follow-up in the first 2 years, followed by a de-intensification after 2 years of follow-up, which will likely result in a lower burden for patients and a better efficiency. However, this follow-up protocol may not be adequate for patients at a higher risk for regrowth, such as patients who have a nCR after 6 months or undergo local excision or contact brachytherapy for a tumour remnant. [22-24] These patients have a higher risk of harboring residual disease in the lumen or regional lymph nodes and should undergo a more intensive follow-up. In less experienced centres physicians might feel more comfortable with a more gradual de-intensification of the current follow-up schedule, as more assessments compensate for missed detections.

ACKNOWLEDGEMENTS

The authors thank the Dutch Watch-and-Wait consortium: Stephanie O. Breukink, M.D., PhD, Maastricht University Medical Center, department of surgery, Maastricht, the Netherlands; Sebastiaan Festen M.D., PhD, OLVG West, department of surgery, Amsterdam, the Netherlands; Eelco J.R. de Graaf, M.D., PhD, IJsselland Hospital, department of surgery, Capelle aan de IJssel, the Netherlands; Brechtje A. Grotenhuis M.D., PhD, The Netherlands Cancer Institute, department of surgery, Amsterdam, The Netherlands; Denise Hilling, M.D., PhD, F.E.B.S., Leiden University Medical Center, department of surgery, Leiden, the Netherlands; Christiaan Hoff, M.D., Medical Center Leeuwarden, department of surgery, Leeuwarden, the Netherlands; Martijn Intven, M.D., PhD, University Medical Center Utrecht, department of radiotherapy, Utrecht, the Netherlands; Niels Komen, M.D., PhD, University

Hospital Antwerpen, department of surgery, Antwerpen, Belgium; Koen CMJ Peeters, M.D., PhD, Leiden University Medical Center, department of surgery, Leiden, the Netherlands; Apollo Pronk, M.D., PhD, F.E.B.S., Diakonessenhuis, department of surgery, Utrecht, the Netherlands; W.H. (Hermien) Schreurs, M.D., PhD, Northwest Clinics, department of surgery, Alkmaar, the Netherlands; Dirk J.A. Sonneveld, M.D., PhD, Dijklander ziekenhuis, department of surgery, Hoorn, the Netherlands; Koen Talsma, M.D., PhD, Deventer Hospital, department of surgery, Deventer, the Netherlands; Jurriaan B. Tuynman, M.D., PhD, Amsterdam University Medical Centres, location VUmc, department of surgery, Amsterdam, the Netherlands; Miranda Kusters, M.D., PhD, Amsterdam University Medical Centres, location VUmc, department of surgery, Amsterdam, the Netherlands; Henderik L. van Westreenen, M.D., PhD, F.E.B.S., Isala, department of surgery, Zwolle, the Netherlands; Johannes H.W. de Wilt, M.D., PhD, RadboudUMC, department of surgery, Nijmegen, the Netherlands; David D.E. Zimmerman, M.D., PhD, F.E.B.S., Elisabeth-Tweesteden Hospital, the Netherlands.

SUPPLEMENTARY MATERIAL

Patient	Location	Modality of detection	PA	Actual LR detection in study cohort in months	LR detection according different schedules in months (difference with actual LR detection in months)				
					Current schedule	Schedule 1	Schedule 2	Schedule 3	Schedule 4
1	Nodal	MRI	ypT0N1	44	48 (4)*	48 (4) *	48 (4) *	48 (4) *	48 (4) *
2	Luminal	Endoscopy	ypT2	20	24 (4) *	21 (1)	20 (0)	24 (4) *	24 (4) *
3	Luminal	Both	ypT2N0	21	24 (3) *	21 (0)	24 (3) *	24 (3) *	24 (3) *
4	Luminal	Endoscopy	ypT1N0	19	24 (5) *	21 (2)	20 (1)	24 (5) *	24 (5) *
5	Luminal	Both	ypT2	19	18 (0)	18 (0)	20 (2)	18 (0)	18 (0)
6	Luminal	Both	ypT1N0	18	18 (0)	18 (0)	20 (2)	18 (0)	18 (0)
7	Luminal	MRI	ypT2	18	18 (0)	18 (0)	20 (2)	18 (0)	18 (0)
8	Luminal	Both	ypT2	17	18 (1)	18 (1)	20 (3) *	18 (1)	18 (1)
9	Luminal	Both	ypT3	16	18 (2)	18 (2)	16 (0)	18 (2)	18 (2)
10	Luminal	Both	ypT2N0	15	18 (3) *	15 (0)	16 (1)	18 (3) *	18 (3) *
11	Luminal	Both	ypT3N0	13	12 (0)	12 (0)	12 (0)	12 (0)	12 (0)
12	Luminal	Both	ypT2N0	12	12 (1)	12 (1)	12 (1)	12 (1)	12 (1)
13	Luminal	Both	ypT3N0	12	12 (0)	12 (0)	12 (0)	12 (0)	12 (0)
14	Nodal	MRI	ypT0N1	12	12 (0)	12 (0)	12 (0)	12 (0)	12 (0)
15	Luminal	Both	ypT3N0	10	9 (0)	9 (0)	9 (0)	9 (0)	12 (2)
16	Luminal	Endoscopy	ypT2N0	10	12 (2)	12 (2)	12 (2)	12 (2)	12 (2)
17	Luminal	Both	ypT2N0	11	12 (1)	12 (1)	12 (1)	12 (1)	12 (1)
18	Luminal	Endoscopy	NA^	10	12 (2)	12 (2)	12 (2)	12 (2)	12 (2)
19	Luminal	Both	ypT2N0	10	12 (2)	12 (2)	12 (2)	12 (2)	12 (2)
20	Luminal	Both	ypT2N0	9	9 (0)	9 (0)	9 (0)	12 (3) *	12 (3) *
21	Luminal	Endoscopy	ypT2N0	9	9 (0)	9 (0)	9 (0)	12 (3) *	12 (3) *
22	Both	Both	ypT3N1	10	9 (0)	9 (0)	9 (0)	12 (2)	12 (2)
23	Luminal	Both	ypT2	9	9 (0)	9 (0)	9 (0)	12 (3) *	12 (3) *
24	Both	Both	ypT2N1c	9	9 (0)	9 (0)	9 (0)	12 (3) *	12 (3) *
25	Luminal	Both	ypT3N0	8	9 (1)	9 (1)	9 (1)	8 (0)	12 (4) *
26	Luminal	Both	ypT2N0	9	9 (0)	9 (0)	9 (0)	8 (0)	12 (3) *
27	Luminal	Both	ypT2N0	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
28	Luminal	Endoscopy	ypT1	7	6 (0)	6 (0)	6 (0)	8 (1)	6 (0)
29	Luminal	Both	NA^^	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)

Patient	Location	Modality of detection	PA	Actual LR detection in study cohort in months	LR detection according different schedules in months (difference with actual LR detection in months)				
30	Luminal	MRI	ypT3N0	7	9 (2)	9 (2)	9 (2)	8 (1)	12 (5) *
31	Luminal	Both	NA^^^	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
32	Luminal	Both	ypT2	7	9 (2)	9 (2)	9 (2)	8 (1)	12 (5) *
33	Luminal	Endoscopy	ypT2	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
34	Luminal	Endoscopy	ypT1	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
35	Luminal	Both	ypT3N0	7	6 (0)	6 (0)	6 (0)	8 (1)	6 (0)
36	Both	Both	ypT3N1	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
37	Luminal	Both	ypT2	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
38	Both	MRI	ypT3N0†	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
39	Luminal	Endoscopy	ypT2N0	4	6 (2)	6 (2)	6 (2)	4 (0)	6 (2)
40	Luminal	Both	ypT2N0	6	6 (1)	6 (1)	6 (1)	8 (2)	6 (1)
41	Both	Both	ypT2N1	6	6 (0)	6 (0)	6 (0)	8 (2)	6 (0)
42	Luminal	Both	ypT2N0	5	6 (1)	6 (1)	6 (1)	8 (3) *	6 (1)
43	Luminal	Endoscopy	ypT1N0	5	6 (1)	6 (1)	6 (1)	8 (3) *	6 (1)
44	Luminal	Both	ypT3N0	5	6 (1)	6 (1)	6 (1)	8 (3) *	6 (1)
45	Nodal	MRI	ypT0N1	4	6 (2)	6 (2)	6 (2)	4 (0)	6 (2)
46	Luminal	Both	NA^^^^	4	6 (2)	6 (2)	6 (2)	4 (1)	6 (2)
47	Luminal	Both	ypT1N0	4	6 (2)	6 (2)	6 (2)	4 (0)	6 (2)
48	Luminal	Both	ypT3N0	4	3 (0)	3 (0)	3 (0)	4 (1)	6 (2)
49	Luminal	Endoscopy	ypT2	3	3 (0)	3 (0)	3 (0)	4 (1)	6 (3) *
50	Luminal	Endoscopy	ypT3	3	3 (0)	3 (0)	3 (0)	4 (1)	6 (3) *

Supplementary table 1 Detailed overview of all patients with a LR and the theoretical difference in detection time-point according to the hypothetical schedules. The actual local regrowth detection was defined as the local regrowth detection according to clinical examinations as actually performed in the study cohort. Delay in local regrowth detection was calculated as the difference between the actual local regrowth detection and local regrowth detection in the different (hypothetical) schedules. Values are in months. *, delay of at least 3 months in detection of local regrowths; In some patients no histology was available, defined as NA, not available; ^, patient underwent total mesorectal excision but no histology available; ^^, patient died before undergoing total mesorectal excision; ^^^, patient refused total mesorectal excision but adenocarcinoma was found in endoscopic biopsy; ^^^^, patient underwent local excision with proven adenocarcinoma in resection specimen; †, patient had a pericolic deposit on MRI due to possible extramural vascular invasion or tumor deposit or pathological lymph node for which he was treated with chemoradiation, chemotherapy and surgery with an on-going pelvic control; LR, local regrowth; FU, follow-up

REFERENCES

1. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
2. Dossa F CTR, Acuna S.A, Baxter N.N. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete respons following neoadjuvant chemoradiation: a systematic review and meta-analysis. *Lancet Gastroenterology heptology*. 2017;(2):501-13.
3. Dattani M, Heald RJ, Goussous G, et al. Oncological and Survival Outcomes in Watch and Wait Patients With a Clinical Complete Response After Neoadjuvant Chemoradiotherapy for Rectal Cancer: A Systematic Review and Pooled Analysis. *Ann Surg*. May 9 2018;268(6):955-67
4. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol*. Nov 2015;22(12):3873-80.
5. van der Sande ME, Figueiredo N, Beets GL. Management and Outcome of Local Regrowths in a Watch-and-wait Prospective Cohort for Complete Responses in Rectal Cancer. *Ann Surg*. Feb 7 2020;doi:10.1097/sla.0000000000003738. Online ahead of print.
6. Martens MH, Maas M, Heijnen LA, et al. Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. *J Natl Cancer Inst*. 2016;108(12):djw171.
7. Maas M, Beets-Tan RG, Lambregts DM, et al. Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *J Clin Oncol*. Dec 10 2011;29(35):4633-40.
8. Richtlijn Colorectaal Carcinoom. 06-07-2020, 2020. Accessed 06-07-2020, 2020. https://www.mdl.nl/sites/www.mdl.nl/files/richlijnen/Richtlijn%20Colorectaal%20Carcinoom.pdf
9. Nahas SC, Rizkallah Nahas CS, Sparapan Marques CF, et al. Pathologic Complete Response in Rectal Cancer: Can We Detect It? Lessons Learned From a Proposed Randomized Trial of Watch-and-Wait Treatment of Rectal Cancer. *Dis Colon Rectum*. Apr 2016;59(4):255-63.
10. Habr-Gama A, Sabbaga J, Gama-Rodrigues J, et al. Watch and wait approach following extended neoadjuvant chemoradiation for distal rectal cancer: are we getting closer to anal cancer management? *Dis Colon Rectum*. Oct 2013;56(10):1109-17.
11. Appelt AL, Pløen J, Harling H, et al. High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. *The Lancet Oncology*. 2015/08/01/ 2015;16(8):919-927.
12. Chino A, Konishi T, Ogura A, et al. Endoscopic criteria to evaluate tumor response of rectal cancer to neoadjuvant chemoradiotherapy using magnifying chromoendoscopy. *Eur J Surg Oncol*. 2018;44(8):1247-53.
13. van der Sommen F, Curvers WL, Nagengast WB. Novel Developments in Endoscopic Mucosal Imaging. *Gastroenterology*. May 2018;154(7):1876-1886.
14. Hupkens BJP, Maas M, Martens MH, et al. Organ Preservation in Rectal Cancer After Chemoradiation: Should We Extend the Observation Period in Patients with a Clinical Near-Complete Response? *Ann Surg Oncol*. Jan 2018;25(1):197-203.
15. Rupinski M, Szczepkowski M, Malinowska M, et al. Watch and wait policy after preoperative radiotherapy for rectal cancer; management of residual lesions that appear clinically benign. *Eur J Surg Oncol*. Feb 2016;42(2):288-96.

16. Renehan AG, Malcomson L, Emsley R, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *The Lancet Oncology*. 2016;17(2):174-183.
17. Sanchez Loria F, Iseas S, O'Connor JM, et al. Non-surgical management of rectal cancer. Series of 68 cases, long follow up in two leading centres in Argentina. *Dig Liver Dis*. Nov 2016;48(11):1372-1377.
18. Colorectaal Carcinoom: Follow-up na chirurgische resectie stadium I-III colon- en rectumcarcinoom. 19/01/2021, 2021. Accessed 19/01/2021, 2021. https://www.oncoline.nl/colorectaalcarcinoom
19. Maas M, Rutten IJ, Nelemans PJ, et al. What is the most accurate whole-body imaging modality for assessment of local and distant recurrent disease in colorectal cancer? A meta-analysis : imaging for recurrent colorectal cancer. *Eur J Nucl Med Mol Imaging*. Aug 2011;38(8):1560-71.
20. Kim DJ, Kim JH, Ryu YH, Jeon TJ, Yu JS, Chung JJ. Nodal staging of rectal cancer: high-resolution pelvic MRI versus ^{18}F-FDGPET/CT. *J Comput Assist Tomogr*. Sep-Oct 2011;35(5):531-4.
21. Hope TA, Kassam Z, Loening A, McNamara MM, Paspulati R. The use of PET/MRI for imaging rectal cancer. *Abdominal radiology (New York)*. Nov 2019;44(11):3559-3568.
22. Sun Myint A, Smith FM, Gollins SW, et al. Dose escalation using contact X-ray brachytherapy (Papillon) for rectal cancer: does it improve the chance of organ preservation? *Br J Radiol*. Dec 2017;90(1080):20170175.
23. Rullier E, Vendrely V, Asselineau J, et al. Organ preservation with chemoradiotherapy plus local excision for rectal cancer: 5-year results of the GRECCAR 2 randomised trial. *Lancet Gastroenterol Hepatol*. May 2020;5(5):465-474.
24. Shaikh I, Askari A, Ourû S, Warusavitarne J, Athanasiou T, Faiz O. Oncological outcomes of local excision compared with radical surgery after neoadjuvant chemoradiotherapy for rectal cancer: a systematic review and meta-analysis. *Int J Colorectal Dis*. Jan 2015;30(1):19-29.

PART II

ONCOLOGICAL OUTCOMES

CHAPTER 5

Is watch-and-wait a safe and effective way to treat rectal cancer in older patients?

Hester E. Haak, Monique Maas, Doenja M. J. Lambregts, Regina G. H. Beets-Tan, Geerard L. Beets, on behalf of the Dutch Watch-and-Wait Consortium

Eur J Surg Oncol. 2020 Mar;46(3):358-362.

ABSTRACT

Introduction

The aim was assess the oncological and functional outcome of the watch-and-wait (W&W) approach in older patients with a clinical (near) complete response after neoadjuvant treatment for rectal cancer.

Material and methods

Patients were included in a W&W-approach (2004-2019) when digital rectal examination, endoscopy and MRI showed a (near) clinical complete response. Patients underwent endoscopy and MRI every 3 months during the first year, and 6-monthly thereafter. Patients aged ≥75 and ≥2 years of follow-up (FU) were selected. Oncological outcomes were assessed with Kaplan-Meier curves. Functional outcome was assessed with colostomy-free rate, Vaizey incontinence score, low anterior resection syndrome-score and International Prostate Syndrome Score.

Results

43/304 (14%) of patients in a W&W-approach met the inclusion criteria. Median FU was 37 (24-109) months. 5/43(12%) developed a local regrowth. All were treated surgically, with one patient experiencing a pelvic failure. Distant metastases occurred in 3/43 patients and 4 patients died, 3 of whom not related to rectal cancer. The 3-year local regrowth-free rate was 88%, 3-year non-regrowth disease-free survival 91%, overall survival 97% and 3-year colostomy-free rate 93%. Overall, the bowel- and urinary dysfunction scores at 3, 12 and 24 months indicated good continence, no or minor LARS and moderate urinary problems.

Conclusion

W&W for older patients with a clinical (near) complete response appears to be a safe alternative to a total mesorectal excision (TME), with a very high pelvic control rate, and few rectal cancer related deaths. Most patients can avoid major surgery and a definitive colostomy, and have a reasonable anorectal and urinary function.

INTRODUCTION

All patients with rectal cancer are faced with important questions regarding treatment options and outcome that are related to anorectal and urogenital dysfunction, the possibility of a permanent stoma, and the balance between quality of life and oncological control. Older patients have additional concerns of operative morbidity and mortality, and the loss of independency that can occur after major rectal cancer surgery. Age by itself is not a good predictor of operative outcome, and there are many tools to assess frailty and the operative risk.[1] At the far end of the spectrum is the very frail patient who clearly cannot tolerate a major rectal resection. However, for the majority of older patients who have some degree of elevated operative risk, major rectal surgery is still an option if the alternative of local control by radiotherapy fails. With a good screening program for frailty and appropriate preoperative, perioperative and postoperative care the mortality rate can be much lower than traditionally estimated. Recent data from a national Dutch registry shows a marked improvement in the last decade, with now a 30-day mortality rate of only 2.4% for ASA III-IV patients aged 71-80, and 4.3% for ASA III-IV patients aged >80.[2]

Despite the improvements in perioperative care, a treatment that avoids major surgery remains of high interest for older patients. The most commonly used alternative treatment option consists of radiotherapy, with or without local excision. The role of radiotherapy and the different treatment schedules and delivery methods have been described in this special EJSO issue by Sun Myint and Gerard.[3] The most optimal outcome of radiotherapy is a complete or a very good response of the tumor with the patient monitored in a watch and wait protocol. The watch and wait protocol for clinically complete responders was championed and further developed by Habr-Gama and caught worldwide attention with the good results published in 2004.[4] The approach initially drew a lot of criticism from the surgical community, but with other series corroborating the initial findings, the watch and wait approach has become a valid alternative to major surgery in patients with a clinical complete response. The data and outcome on the largest series of 880 patients was presented by the International Watch and Wait Data registry in 2018.[5] It showed a regrowth rate of 25%, with the majority located in the bowel wall and occurring in the first two years. The 5-year overall survival was 85%, with many patients dying from unrelated causes, corresponding with a disease specific survival of 97%. The details of the treatment of the local regrowth were not always available, and it was estimated that the rate of locally unsalvageable disease was 1% at most.

In The Netherlands there is an ongoing collaborative network since 2004 that focusses on the watch and wait strategy (W&W), with a database containing detailed information and studies on imaging, oncological outcome and quality of life.[6-9] The aim of the present study is to provide information on the outcome of a watch and wait policy for rectal cancer in older patients in this collaborative Dutch database.

MATERIALS AND METHODS

Patients

Rectal cancer patients who were diagnosed with a clinical complete response after neoadjuvant treatment between 2004 and 2019 were offered a W&W program through the national network registration. The majority of patients were prospectively included with an informed consent in two IRB approved studies, registered in clinicaltrials.gov since 2009 (NCT00939666 and NCT02278653). Other patients were retrospectively included in the registry, for whom informed consent was waived by the local institutional review board. Patients who were offered and opted for a W&W were aware that this approach was an alternative treatment that deviated from the standard guideline treatment of TME resection. Inclusion criteria for the present study and analysis were: 1) biopsy proven primary rectal adenocarcinoma without distant metastasis at baseline, 2) neoadjuvant treatment with long course CRT (28x1.8Gy with 2x825 mg/m^3 capecitabine) or short course radiotherapy (5x5Gy) and a prolonged waiting interval 3) minimum age of 75 years or older, 4) minimum of 2 years follow-up in a W&W program, and 5) (near) complete response at first or second response evaluation.

Response assessment

The response was evaluated 8-12 weeks after the end of radiotherapy with digital rectal examination, flexible sigmoidoscopy and standard MRI with additional diffusion weighted imaging (MRI-DWI). Criteria for a clinical complete response have been described previously, and consist of no lesions felt at digital rectal examination (DRE), a typical white scar with or without telangiectasia at endoscopy and no signs of residual disease on MRI.[6] Patients with a very good but not a complete response were labelled 'near-complete responses', consisting of superficial ulceration or irregular persisting erythema on endoscopy and intermediate/low residual signal on T2W-MRI and/or small foci of diffusion restriction on MRI-DWI. These patients were offered a second reassessment, 8-12 weeks after the first evaluation, and could be included in the study.

Follow-up

The standard follow-up for rectal cancer consisted of computed tomography scan (CT) and carcinoembryogenic (CEA) antigen measurements as recommended by national guidelines every year for 5 years. W&W patients were additionally followed with a DRE, endoscopy and MRI (+DWI) every 3 months the first year, and 6-monthly thereafter up to 5 years.

Functional outcomes

As a part of the second prospective study (NCT02278653) pelvic functional outcomes were assessed with questionnaires. Defecation problems were assessed with the Vaizey score and the low anterior resection syndrome (LARS) score.[10, 11] The Vaizey score is a faecal incontinence score based on defecation pattern of the previous 4 weeks and consists of questions regarding frequency, consistency of stools lost and effect on lifestyle. The range

of the score is 0-24 in which a score of 12 or higher is indicated with major incontinence. The LARS-score is a score intended for patients who underwent low anterior resection for rectal cancer, and consists of 5 questions regarding frequency, clustering, urgency and incontinence for flatus or liquid stools. The range of this score is 0-42 and is divided into no, minor or major LARS (0-20 points; 21-29 points; 30-42 points respectively). Colostomy-free rate served also as a measure of quality of life. Urinary problems were assessed with the International Prostate Symptom Score (IPSS).[12] Although this score was designed to assess bladder function in patients with benign prostate hypertrophy, it was judged to be the most suited for the purpose when no other and more focussed questionnaires were available. It consists of 7 questions addressing frequency, urgency, intermittency, weak stream, nocturnia, straining, incomplete bladder emptying and quality of life. The IPSS ranges from 0-35 and is divided into mild (0-7), moderate (8-19) and severe symptoms (20-35).

Statistical analysis

Baseline characteristics were provided. Local regrowth free rate (LRFR), colostomy-free rate (CFR), non-regrowth disease-free survival (NRDFS) and overall survival (OS) were estimated with Kaplan-Meier curves. Duration of follow-up was calculated between end of last CRT and time to event of interest or last follow-up date. Local regrowth was defined as luminal regrowth or involvement of loco regional lymph nodes and local control was defined as the absence of local regrowth. NRDFS was defined as the absence of pelvic failure (local recurrence after delayed TME for local regrowth), the absence of distant metastasis or absence of death. Overall survival was defined as the absence of death. Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS version 25.0, Inc., Chicago, IL).

RESULTS

Patient characteristics

Between 2004-2019 a total of 304 rectal cancer patients followed a W&W approach with a minimum follow-up of two years. 43/304(14%) patients were 75 years or older and constitute the group of patients for the present study. The majority of these patients were included prospectively (n=27). Median age was 78 years (range 75-87) and 29/43 (67%) patients were male. The majority of patients (40/43) received neoadjuvant CRT, and 7/43 (17%) of patients received adjuvant chemotherapy after CRT. Nineteen of 43 patients (44%) were included immediately 8-12 weeks after CRT and the remaining 24/43 (56%) patients were included at reassessment 8-12 weeks after the first evaluation. Median FU was 37 (24-109) months. See *Table* 1 for detailed baseline characteristics.

	Total cohort (N=43)
Age, median (range), years	78 (75-87)
Sex (male)	67% (29/43)
Clinical T stage	
T1	2% (1/43)
T2	16% (7/43)
T3	67% (29/43)
T4	14% (6/43)
Clinical N stage (N+)	61% (26/43)
Distance anal verge (cm)	
<5	71% (31/43)
>5	29% (12/43)
Adjuvant chemotherapy	17% (7/43)
Neo-adjuvant therapy	
CRT	93% (40/43)
5x5 Gy	5% (2/43)
Other	2% (1/43)
Timing inclusion	
Immediate	44% (19/43)
Reassessment	56% (24/43)
Follow-up time, median (range), months	37 (24-109)

Table 1 Baseline characteristics of all patients. Data are % (n/N) unless otherwise stated. CRT = chemoradiation, Gy = Gray

Oncological outcomes

Five of 43 patients (12%) developed a local regrowth, all of which were luminal and within two years (range 5-18). Four were detected both on MRI and endoscopy and one was diagnosed only with endoscopy. Three patients underwent a TME resection, one low anterior resection (LAR) and 2 abdominoperineal resection (APR). The remaining two patients underwent local excision. One patient had a local recurrence after TME, leading to an overall pelvic control in 42/43 patients (98%).

The local regrowth free rate, non-regrowth disease-free survival and overall survival at 3-years was 88%, 91% and 97% respectively (see Figure 1). Three patients developed distant metastasis (1 lung; 1 lung and liver; 1 peritoneal metastasis) and as mentioned above there was one patient with a local recurrence. Four patients eventually died of the following causes: metastatic disease (one patient), other cancer with no evidence or rectal cancer recurrence (2 patients), and a ruptured suprarenal aneurysm (one patient).

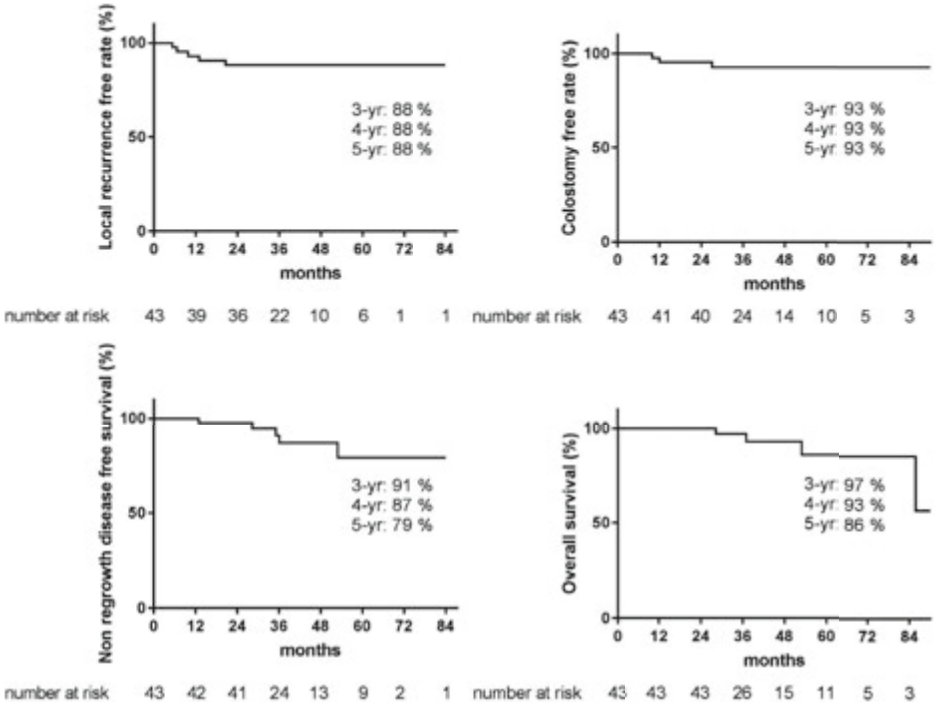


Figure 1 Kaplan Meier survival curves (local regrowth free rate, non-regrowth disease free survival, overall survival and colostomy-free rate) for all patients

Functional outcomes

Only three of 43 patients ended up with a stoma (2 after APR for a regrowth, and one after LAR due to persistent incontinence), resulting in a 3-year colostomy-free rate of 93% (see Figure 1). Thirteen patients were included in the prospective study mainly focussing on functional problems (NTC02278653), of whom 9/13 (69%) completed the Vaizey score and LARS-score, and 7/13 (54%) completed the IPSS. All responders had a sustained complete response during follow-up.

The mean Vaizey score after 3, 12 and 24 months was 4.4 (SD 3.5), 4.2 (SD 4.2) and 5.4 (SD 5.1) respectively, indicating a good continence. Two out of 7 patients (29%) had major incontinence (score ≥ 12) at 24 months. The mean LARS-score at 3, 12 and 24 months was 19.7 (SD 11.1), 19.6 (SD 12.3) and 21.3 (SD 9.4) respectively. Two patients had major LARS (score 30-42) during all time-points (3, 12 and 24 months). See Figure 2 for detailed results. The mean IPSS was 7.3 (SD 6.5) at 3 months, 9.1 (SD 5.9) at 12 months and 9 (SD 5) at 24 months. Only one (14%) of 7 patients had severe urinary problems after 3 months. See Figure 3 for detailed results.

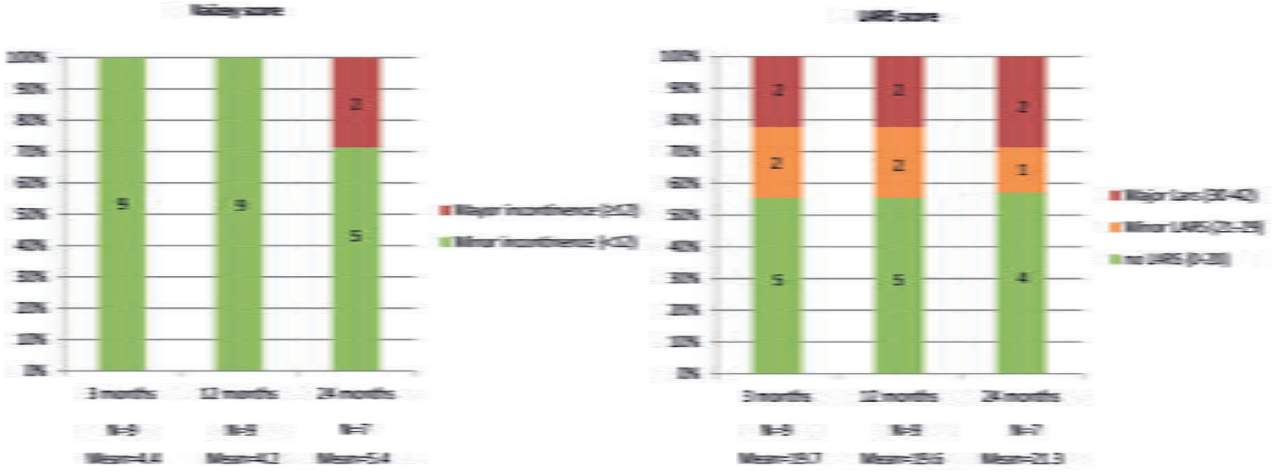


Figure 2 Defecation problems according the Vaizey- and low anterior resection syndrome (LARS) -score.

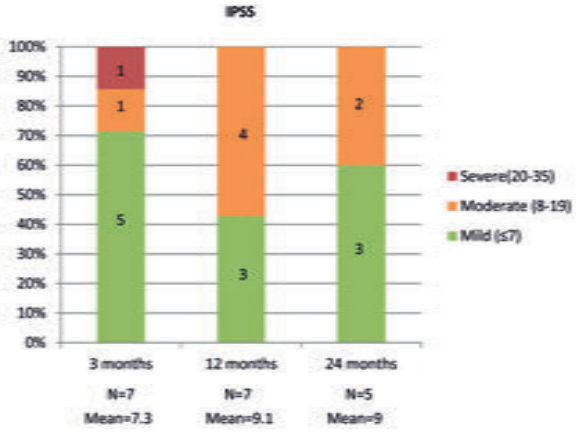


Figure 3 Urinary problems according International Prostate Symptom Score (IPSS)

DISCUSSION

The outcome of older patients with a complete or near complete response after radiotherapy for rectal cancer who are followed in a watch and wait protocol in the Dutch national cohort series is very good. The 3-year overall survival is 97%, with death mostly due to other causes. The 3-year local regrowth rate is 12%, with delayed surgery resulting in a high pelvic control rate of 98%. There was a 3-year colostomy-free rate of 93%, and although the sample size of the functional outcome assessment was too small for a reliable analysis, the results suggest a reasonably good anorectal and urinary function.

The overall and disease specific survival is similar to that described in the largest series of 880 patients in the International Watch and Wait Data registry in 2018.[5] The 25% 2-year regrowth rate in that series was much higher than in the current study. The baseline tumour characteristics in the two studies were similar, with the majority a T3 tumour with positive nodes, consistent with the accepted indication for neoadjuvant radiotherapy. The difference in regrowth rate most likely reflects the difference in the selection criteria for a complete response, and whether or not a second reassessment is performed for 'near complete responses' before taking a decision. The Dutch collaborative network generally uses a strict definition of complete response, and is liberal with second assessments before deciding on taking patients to surgery or start a W&W approach, leading to the relatively low regrowth rate when compared to other centres. There is a general concern that omitting TME surgery exposes patients to an oncological risk by uncontrolled pelvic disease and metastatic disease originating from the regrowth. The IWWD data suggest that with a good follow up program the risk for uncontrolled pelvic disease after W&W for complete responders is 1% at most, in line with the current findings. The excess metastatic risk in W&W is more difficult to estimate but the very high disease specific 5-year survival of 95% suggests that the risk of metastases is more related to tumour biology than to the omission of immediate surgery.[5] For the frail and older patient this oncological risk is more easily counterbalanced by an increased operative risk and a decreased life expectancy than for a younger patient. With a decision-analytic model Smith et al. even suggested an improved survival for W&W compared to TME surgery for increasing age and comorbidity.[13]

The patients in the current study were 75 or older, but the fact that all patients with regrowth were treated surgically and the high overall survival suggests that the majority of patients was reasonable fit, and underwent neoadjuvant radiotherapy with the intent to have a TME resection. The very old and frail patients who cannot tolerate major rectal surgery and are treated with radiotherapy as a definitive treatment are underrepresented in the present study, as patients are only registered when they have a complete clinical response at the time of response assessment. Standard neoadjuvant radiotherapy results in complete response in only 15-30%, depending on the size and T-stage of the tumour.[14] When the explicit goal is to obtain higher local control rates with radiotherapy there are a number of options to deliver a higher dose, with internal or external boost techniques, as reviewed

by Sun Myint et al. and Bujko et al.[3, 15] Another approach that is expected to yield higher response rates is to add systemic therapy either as induction or consolidation therapy in combination with radiotherapy, but the added toxicity will limit the use in older patients.[16] Major rectal surgery can also be avoided by performing a transanal local excision of a small tumour remnant when there was a good response after radiotherapy.[17]

In addition to the benefit of avoiding postoperative complications and mortality major surgery, older patients have a high interest in maintaining a high quality of life and remaining as independent from caretakers as possible. In the current study 71% of patients presented with a tumour within 5cm form the anal verge. For most patients, standard TME surgery would have resulted in a permanent colostomy, or a poor function in a very low anastomosis after neoadjuvant therapy.[18] The 3-year colostomy-free rate was 93%, with only the three patients requiring delayed TME surgery resulting in a permanent colostomy. This is in line with the 95% reported by Martens et al. and the crude overall 90% by Smith et al.[6, 19] The paper on the UK OnCoRe project by Renehan et al. describes a lower 74% 3-year colostomy-free survival, related to the broader inclusion and the resulting higher 38% 3-year local regrowth rate.[20] Radiotherapy by itself has some negative effect on anorectal function, and up to one third of patients in a W&W program are reporting major LARS symptoms.[9] Although in the current study the number of available questionnaires was too low for a reliable analysis, the results suggest a reasonably good anorectal and urinary function in the majority of patients. The most commonly reported symptoms are clustering and urgency, and in our experience almost all patients can handle this without much interference with daily activities. The majority of patients are very satisfied to have avoided the short and long-term side effects of major surgery and a permanent colostomy.

There are some limitations of the present study. Because of the design of the registry there is an underrepresentation of very old and frail patients who are treated with radiotherapy as a definitive treatment. The encouraging outcome of the W&W approach in clinical complete responses in reasonable fit older patients in the current study can in our view be extrapolated to the more frail patients, who have even more to gain from avoiding major surgery. A further limitation is the very small number of data on quality of life. Given the importance of this outcome measure in this group of patients, more data should be generated in preferably prospective studies.

CONCLUSION

The outcome of older patients with a complete or near complete response after radiotherapy for rectal cancer who are followed in a watch and wait protocol is very good. The pelvic control rate is very high, and the majority of deaths are not related to rectal cancer. Most patients can avoid major surgery and a definitive colostomy, and have a reasonable anorectal and urinary function.

ACKNOWLEDGEMENTS

The authors thank the Dutch Watch-and-Wait consortium collaborators**:** J. (Jarno) Melenhorst, M.D., PhD, Maastricht University Medical Center, department of surgery, Maastricht, the Netherlands; M.E. (Marit) van der Sande, M.D., The Netherlands Cancer Institute, department of surgery, Amsterdam, The Netherlands; H.L. (Erik) van Westreenen, M.D., PhD, Isala, department of surgery, Zwolle, the Netherlands; A.K. (Koen) Talsma, M.D., PhD, Deventer Hospital, department of surgery, Deventer, the Netherlands; S.O. (Stephanie) Breukink, M.D., PhD, Maastricht University Medical Center, department of surgery, Maastricht, the Netherlands; D.D.E. (David) Zimmerman, M.D., PhD, Elisabeth TweeSteden Hospital, department of surgery, Tilburg, the Netherlands; D. (Denise) Hilling, M.D., PhD, Leiden University Medical Center, department of surgery, Leiden, the Netherlands; J.H.W. (Hans) de Wilt, M.D., PhD; RadboudUMC, department of surgery, Nijmegen, the Netherlands; KCMJ (Koen) Peeters, M.D., PhD, Leiden University Medical Center, department of surgery, Leiden, the Netherlands; E.J.R. (Eelco) de Graaf, M.D., PhD, IJsselland Hospital, department of surgery, Capelle aan de IJssel, the Netherlands; J.B. (Jurriaan) Tuynman, M.D., PhD, Amsterdam UMC, Vrije Universiteit Amsterdam, department of surgery, Amsterdam, the Netherlands; D.J.A. (Eric) Sonneveld, M.D., PhD, Westfriesgasthuis, department of surgery, Hoorn, the Netherlands; N. (Niels) Komen, M.D., PhD, Universitair Ziekenhuis Antwerpen, department of surgery, Antwerpen, Belgium. A (Apollo) Pronk, M.D., PhD, Diakonessenhuis, department of Surgery, Utrecht, the Netherlands; M (Martijn) Intven, M.D., Ph, University Medical Center Utrecht, department of radiotherapy, Utrecht, the Netherlands; W.H. (Hermien) Schreurs, M.D., PhD, Noordwest Ziekenhuisgroep, department of surgery, Alkmaar, the Netherlands; C. (Christiaan) Hoff, M.D., Medical Center Leeuwarden, department of surgery, Leeuwarden, the Netherlands.

REFERENCES

1. Rostoft S, Hamaker ME. Basic geriatric principles for colorectal surgeons: How to optimize assessment and care of older patients in the perioperative period. *Eur J Surg Oncol*. 2020;46(3):310-315.
2. de Neree Tot Babberich MPM, Detering R, Dekker JWT, et al. Achievements in colorectal cancer care during 8 years of auditing in The Netherlands. *Eur J Surg Oncol*. Sep 2018;44(9):1361-1370.
3. Sun Myint A, Gerard JP. Role of radiotherapy in the treatment of rectal cancer in older patients. *Eur J Surg Oncol*. Mar 2020;46(3):349-357.
4. Habr-Gama A, Perez RO, Nadalin W, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg*. Oct 2004;240(4):711-7; discussion 717-8.
5. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
6. Martens MH, Maas M, Heijnen LA, et al. Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. *J Natl Cancer Inst*. Dec 2016;108(12); djw171.
7. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol*. Nov 2015;22(12):3873-80.
8. Lambregts DMJ, Maas M, Boellaard TN, et al. Long-term imaging characteristics of clinical complete responders during watch-and-wait for rectal cancer-an evaluation of over 1500 MRIs. *Eur Radiol*. Aug 2019;30:272-280.
9. Hupkens BJP, Martens MH, Stoot JH, et al. Quality of Life in Rectal Cancer Patients After Chemoradiation: Watch-and-Wait Policy Versus Standard Resection - A Matched-Controlled Study. *Dis Colon Rectum*. Oct 2017;60(10):1032-1040.
10. Vaizey CJ, Carapeti E, Cahill JA, Kamm MA. Prospective comparison of faecal incontinence grading systems. *Gut*. Jan 1999;44(1):77-80.
11. Emmertsen KJ, Laurberg S. Low anterior resection syndrome score: development and validation of a symptom-based scoring system for bowel dysfunction after low anterior resection for rectal cancer. *Ann Surg*. May 2012;255(5):922-8.
12. Barry MJ, Fowler FJ, Jr., O'Leary MP, et al. The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol*. Nov 1992;148(5):1549-57; discussion 1564.
13. Smith FM, Rao C, Oliva Perez R, et al. Avoiding radical surgery improves early survival in elderly patients with rectal cancer, demonstrating complete clinical response after neoadjuvant therapy: results of a decision-analytic model. *Dis Colon Rectum*. Feb 2015;58(2):159-71.
14. Maas M, Nelemans PJ, Valentini V, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol*. Sep 2010;11(9):835-44.
15. Bujko K, Glynne-Jones R, Papamichael D, Rutten HJT. Optimal management of localized rectal cancer in older patients. *J Geriatr Oncol*. Nov 2018;9(6):696-704.

16. 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*. Jul 2019;271(3):440-448.
17. Arezzo A, Lo Secco G, Passera R, et al. Individual participant data pooled-analysis of risk factors for recurrence after neoadjuvant radiotherapy and transanal local excision of rectal cancer: the PARTTLE study. *Tech Coloproctol*. Aug 2019;23:831-849.
18. Battersby NJ, Bouliotis G, Emmertsen KJ, et al. Development and external validation of a nomogram and online tool to predict bowel dysfunction following restorative rectal cancer resection: the POLARS score. *Gut*. Apr 2018;67(4):688-696.
19. Smith JJ, Strombom P, Chow OS, et al. Assessment of a Watch-and-Wait Strategy for Rectal Cancer in Patients With a Complete Response After Neoadjuvant Therapy. *JAMA oncology*. Jan 10 2019:e185896.
20. Renehan AG, Malcomson L, Emsley R, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *Lancet Oncol*. Feb 2016;17(2):174-83.

CHAPTER 6

Time pattern to metastases after chemoradiation and watch-and-wait compared to chemoradiation and surgery for rectal cancer.

Hester E. Haas, Monique M. Maas, Renaud L.M. Tissier, Rodrigo O. Perez, Angelita Habr-Gama, Cornelis J. H. van de Velde, Andrew G. Renehan, Geerard L. Beets, on behalf of the International Watch-and-Wait Consortium.

In preparation

PART III

RECENT ADVANCED IMAGING TECHNOLOGIES

CHAPTER 7

Modern MR imaging technology in rectal cancer; there is more than meets the eye

Hester E. Haak, Monique Maas, Stefano Trebeschi, Regina G. H. Beets-Tan

Front Oncol. 2020 Oct 8;10:537532.

ABSTRACT

MR imaging (MRI) is now part of the standard work up of patients with rectal cancer. Restaging MRI has been traditionally used to plan the surgical approach. Its role has recently increased and been adopted as a valuable tool to assist the clinical selection of clinical (near) complete responders for organ preserving treatment. Recently several studies have addressed new imaging biomarkers that combined with morphological provides a comprehensive picture of the tumor. Diffusion weighted MRI (DWI-MRI) has entered the clinics and proven useful for response assessment after chemoradiotherapy. Other functional (quantitative) MRI technologies are on the horizon including artificial intelligence modeling. This narrative review provides an overview of recent advances in rectal cancer (re)staging by imaging with a specific focus on response prediction and evaluation of neoadjuvant treatment response. Furthermore, directions are given for future research.

INTRODUCTION

In the past, rectal cancer surgery was associated with a local recurrence rate of up to 30%.[1] Due to the introduction of total mesorectal excision (TME) in combination with neoadjuvant (chemo-) radiotherapy and optimal staging by MR imaging (MRI), this rate is now down to less than 3%.[2] MR-imaging plays a pivotal role in primary staging to stratify the patient to the right treatment according to the risk for local recurrence. Restaging MRI after neoadjuvant treatment can accurately show downsizing and downstaging of the tumor and plan the surgical approach. In approximately 15-20% of the patients neoadjuvant chemoradiation (CRT) leads to (near) complete response. Organ preserving approach such as a watch-and-wait (W&W) has shown to be a safe alternative for surgery provided that the patient is managed in expert centers, equipped with modern imaging and endoscopic technology and a dedicated multidisciplinary team. To be able to safely apply W&W, selection of patients with a (near) complete response is key. Functional MRI techniques capture changes in tumor perfusion and microstructure before morphological changes become apparent.[3] A well-established functional MRI is diffusion-weighted MRI (DWI) which analyses the movement or "diffusion" of water molecules in which tissues with high cellularity (i.e. tumors and lymph nodes) have restricted diffusion (high signal), while normal tissue and fibrosis will lead to free diffusion (low signal). DWI is now part of the standard restaging MRI. In addition, other techniques such as perfusion MRI [dynamic contrast enhanced (DCE) MRI] and artificial intelligence modeling are being explored. This narrative review provides an overview of recent advances in rectal cancer imaging with a specific focus on response prediction and evaluation of neoadjuvant treatment response. Furthermore, directions are given for future research.

Baseline imaging

MRI and endorectal ultrasound (EUS) are the established modalities for rectal cancer imaging. MRI is the most accurate modality to assess the tumor extent, nodal involvement and guide treatment planning. For the distinction between T1 and T2 tumors, MRI is not accurate, therefore EUS is used for this specific purpose.[4] However, morphological imaging does not provide information on tumor biology. Martens et al. reviewed the available literature on different volumetric methods and showed that whole-volume measurements on pre-CRT imaging reaches the highest accuracy of 71-73% for prediction of response to CRT.[5] Studies which focused on both MR and DWI volumetry at baseline, MRI showed low to moderate performances for predicting the response to CRT as compared to volumetric changes at post-CRT imaging (AUC of 0.57-0.73% vs. 0.63-0.77).[6, 7] The largest evidence is for DWI. A low pre-CRT apparent diffusion coefficient (ADC) at baseline DWI has by several studies shown to be significantly related with pathological complete response (pCR) and good response.[6, 8-11] A possible explanation is that tumors with a high ADC value have more areas of necrosis which makes them less sensitive for radio- and chemotherapy.[12] Despite some initial promising results for ADC subsequent literature showed conflicting results and considerable variability in reported cut off ADC values.[13] Hence ADC measurements has

not gained a significant role in response evaluation of rectal cancer treatment. Intravoxel incoherent motion (IVIM) uses low-b-values of DW-images to extract the perfusion fraction hence providing information on the tumor microvasculature without administration of intravenous contrast. Several studies show that IVIM is promising in prediction of response.[14, 15] Diffusion kurtosis uses very high b-values and reflects intratumoral heterogeneity. The first few studies have not shown superiority of kurtosis imaging above ADC parameters in predicting the treatment response.[16, 17] Two studies did show some potential for kurtosis imaging to predict response.[18, 19] However, due to the lack of standardization and lack of strong evidence both IVIM and kurtosis imaging are currently only explored in a research setting.[20]

DCE-MRI is a method which measures the inflow of intravenously injected contrast agents into the tumor and leakage of contrast into the extracellular space on T1W-MRI and provides direct information of the tissue perfusion. DCE-MRI can be analyzed quantitatively (by measurement of the perfusion of a voxel-by-voxel basis) or by semiquantitative analyses (in which a signal intensity time curve is plotted to assess parameters such as time to peak enhancement or area under the curve). DCE-MRI can provide valuable information on tumor biology (aggressiveness and the degree of angiogenesis) and initial studies have shown promise in the prediction of response.[19, 21-25] Several studies showed that patients who achieved a pCR had significantly higher perfusion parameters [Ktrans, Kep (volume of extracellular space), and Ve (constant of flow rate)] than those who did not.[21, 22, 25] Another group showed that the 'late slope' of the signal enhancement curve after administration of contrast on baseline DCE-MRI was able to differentiate between good and poor responders with an AUC of 0.90[23], although this study used a macromolecular blood pool contrast agent 'gadofosveset' instead of the in clinics routinely applied micromolecular contrast 'Gadolinium DTPA'. So far, DCE-MRI has not found its way to clinical practice due to the relatively high intra- and inter-tumor variation, need for intravenous contrast agents and lack of robustness of the technique. Research for optimization as well as standardization of the technique is much needed.[4, 26] Table 1 provides an overview of the accuracy and predictive values of the different MRI techniques in a primary setting before neoadjuvant treatment.

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC
Baseline						
T2W volumetry	31-55 [5-7]	74-83 [5-7]	31-50 [5-7]	79-88 [5-7]	71-73 [5-7]	0.57-0.73 [6,7]
DWI volumetry	57-65 [6,7]	76-78 [6,7]	37-50 [6,7]	82-91 [6,7]	72-74 [6,7]	0.63-0.77 [6,7]
ADC	38-69 [6,8,11,16]	68-86 [6,8,11,16]	35-42 [6,8,11,16]	78-91 [6,8,11,16]	66-81 [6,8,11,16]	0.55-0.77 [6,8,16,19]
IVIM						
Pseudo-diffusion coefficient	60 [14]	84 [14]	55 [14]	87 [14]	NA	0.74 [14]
Perfusion fraction	80 [14]	72 [14]	47 [14]	92 [14]	NA	0.71 [14]
Kurtosis						
Mean kurtosis coefficient	93-100 [16,18,19]	67-81 [16,18,19]	62 [16]	97 [16]	84 [16]	0.86-0.91 [16,18,19]
Mean diffusion coefficient	64 [16]	62 [16]	36 [16]	84 [16]	63 [16]	0.56 [16]
DCE-MRI						
Kep	100 [25]	70 [25]	NA	NA	NA	0.67-0.75 [21,19,25]
Ktrans	75-100 [21,25]	60-73 [21,25]	NA	NA	NA	0.57-0.92 [21,19,25]
Ve	83 [19]	83 [19]	NA	NA	NA	0.66-0.86 [21,19]
Radiomics	80-96 [56,59]	67-89 [56,59]	NA	NA	93-95 [56]	0.66-0.97 [56-59,63]
Restaging						
T2W volumetry	14-100 [5-7, 30]	30-97 [5-7, 30]	26-82 [5-7, 30]	82-100 [5-7, 30]	44-86 [5-7, 30]	0.70-0.82 [6,7, 30]
DWI volumetry	70-79 [6,7,30]	95-100 [6,7, 30]	88-100 [6,7, 30]	80-94 [6,7, 30]	84-94 [6,7, 30]	0.91-0.93 [6,7, 30]
T2W-MRI fibrotic patterns	94 [31]	77 [31]	88 [31]	87 [31]	88 [31]	
ADC	46-93 [6,11,16, 30]	56-81 [6,11,16, 30]	27-71 [6,11,16, 30]	65-97 [6,11,16, 30]	53-78 [6,11,16, 30]	0.54-0.82 [6,7, 30]
IVIM						
True diffusion coefficient	60 [14]	97 [14]	86 [14]	89 [14]	NA	0.80 [14]
Kurtosis						
Mean kurtosis coefficient	93 [16]	83 [16]	65 [16]	97 [16]	86 [16]	0.91 [16]
Mean diffusion coefficient	93 [16]	71 [16]	52 [16]	97 [16]	73 [16]	0.87 [16]
DCE-MRI						
Ktrans	83-100 [22]	67-78 [22]	29-33 [22]	97-100 [22]	NA	0.81-0.84 [22]
MT imaging	88 [41]	90 [41]	70 [41]	96 [41]	NA	0.96 [41]
3-modality approach	71 [42]	97 [42]	NA	NA	NA	0.89 [42]
Radiomics	100 [61]	91 [61]	72-90 [55,61]	95-100 [55, 61]	94 [55]	0.93-0.98 [55,61]

Table 1 Overview of the accuracy and predictive values of the different MRI techniques during baseline and restaging to identify pathological complete response. The table provides an overview of studies who evaluated MR techniques during baseline or restaging to identify a pathological complete response. Studies who evaluated MR techniques to identify good responders vs. non-responders or parameters such as change values determined pre- and post-CRT were not selected. PPV=positive predictive value; NPV=negative predictive value; AUC=area under the curve receiver operating characteristics; T2W=T2-weigthed; DWI=diffusion-weighted MRI; ADC=apparent diffusion coefficient; IVIM=intravoxel incoherent motion; DCE-MRI=dynamic contrast enhanced MRI; Kep =rate constant; Ktrans=volume transfer constant; Ve=extravascular space; MT imaging=magnetization transfer imaging

Response evaluation

As for primary staging, morphological MRI is also the main imaging modality to evaluate the luminal response and to identify extraluminal findings or remaining malignant nodes after neoadjuvant treatment. However, standard T2-weighted MRI lacks the ability to accurately evaluate response to neoadjuvant therapy because of the difficult distinction between fibrosis with and without viable tumor. Reported accuracies are 43-60% for detection of residual tumor after CRT.[27-29] Studies addressing T2W and DWI volumetry have shown that the decrease in volume and absolute volume after CRT were correlated with response T2W volumetry and showed accuracies over 80% for the assessment of complete responders.[6, 7] DWI-volumetry using whole tumor volume manual delineation outperformed T2W volumetry (AUCs up to 0.93) in several studies.[6, 7, 30] However, manual volumetry is time-consuming making it less useful to apply in clinical practice.[13] Lambregts et al. proposed a method to qualitatively assess the fibrotic pattern that appears after CRT and showed that the exact type of fibrotic pattern on restaging T2W-MRI helps to evaluate the response after CRT.[31] They found that the fibrotic pattern follows the pattern of the primary tumor. For example, a polypoid or (semi)circular tumor shows a sharply demarcated semicircular fibrotic wall after CRT, and an irregular or spiculated tumor often shows irregular fibrotic thickening of the wall on restaging MRI. Only 25% of the patterns were easy to interpret i.e. a normalized wall in the complete responders and bulky tumoral mass in the patients with residual disease. In the majority of the irradiated patients, however, different fibrotic patterns were seen (i.e. a mixed signal or irregular aspect on T2) which were difficult to interpret.[31, 32] The magnetic resonance tumor regression grade (mrTRG), adapted from a similar TRG classification used in histopathology[33] categorized response into a scale from TRG-1 (only fibrosis, probably complete response) to TRG-5 (no fibrosis, probably residual disease)[34, 35]. Siddiqui et al showed that this metric has a good interobserver agreement and in 90% the radiologists correctly identified poor responders. However, in only 66% of the cases the radiologists correctly selected good responders.[35] An on-going randomized controlled trial aims to assess the ability of mrTRG to direct management.[36] The results of this trial will show whether mrTRG based stratification will impact outcome.

The value of functional MR parameters for response evaluation after CRT have been explored. Several studies showed that a higher value of ADC and a larger increase in ADC are both associated with a good response to CRT.[10, 37, 38] A meta-analysis described pooled sensitivities and specificities of 68% and 69% for pretreatment ADC for the prediction of pCR after CRT, and of 72 and 78% respectively for the increase in ADC after CRT.[11] One study showed that an increase of an IVIM coefficient was seen after CRT with a significantly higher value in good versus poor responders.[39] Another study showed that diffusion kurtosis imaging was feasible to assess response and superior to mrTRG.[17] However, both techniques are far from ready to be implemented in clinics and remain in research setting.[16, 17, 39]

A repeated finding in multiple DCE MRI studies is that a large decrease of Ktrans after CRT is predictive for (complete) response.[26] For most other (semi-) quantitative DCE parameters

after CRT no robust conclusions can be drawn[26] which is the reason why DCE-MRI is not routinely applied in clinical practice. A less studied functional imaging technique is magnetization transfer (MT) imaging (traditionally applied in brain imaging). MT imaging explores differences in the magnetization interaction. The transfer of magnetization (MT ratio) between protons bound to macromolecules (collagen) and free/unbound water protons is high in case of collagen rich tissue (fibrosis) and may be useful to discriminate residual disease from post-CRT fibrosis.[40, 41] Yet evidence is limited and MT imaging is only explored in research setting.

The highest accuracy to identify complete responders after CRT has been reached with the use of a three-modality approach[42] including digital rectal exam (DRE) with endoscopy, T2W-MRI and DWI. Maas et al described that when the combination of DRE, endoscopy, T2W-MRI and DWI all indicate a complete response, this diagnosis is correct in 98% of the patients[42] Figure 1. Endoscopy was shown to be an invaluable tool for response evaluation, with MRI being an important adjunct to assess the extramural parts of the tumor and nodes. This method has been adopted globally in the selection of patients for organ preservation.[43] Despite the good results, up to 15% of complete responders are still missed, due to the fact that many complete responding patients may show some findings that are often associated with residual tumor (e.g. ulcers at endoscopy, focal diffusion signal on DWI, irregular nodes on T2W MRI among others).[44-46] Unfortunately, biopsies are of limited clinical value in this setting, because of the risk for sampling error and a risk for false positive findings (e.g. dysplasia in biopsy, but complete response in the TME specimen).[44, 46-48] An overview of the accuracy and predictive values of the different MRI techniques during restaging is given in Table 1.

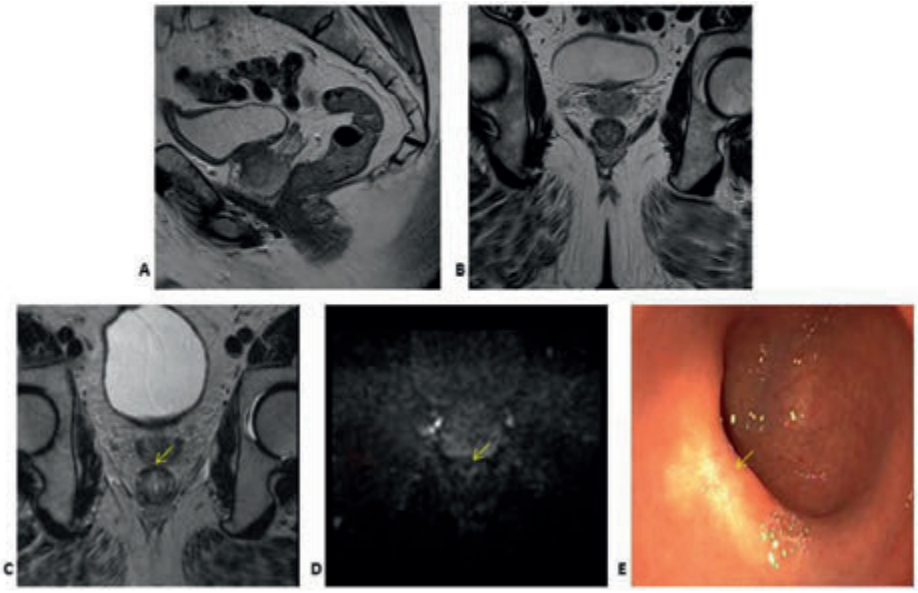


Figure 1 Three-modality approach with combination of T2W-MRI, DWI and endoscopy of the selection of a patient with a clinical complete response. A low rectal tumor is seen on MRI (A,B). On restaging MRI 8 weeks after completion of chemoradiotherapy (C) only minimal fibrosis (yellow arrow) is seen anteriorly in the rectal wall. On restaging DWI (D) there is absence of high diffusion signal (yellow arrow). Clinical assessment by endoscopy (E) reveals a white scar with telangiectasia (yellow arrow).

Artificial intelligence

Artificial intelligence models hold promise in cancer imaging. These approaches aim to use computer algorithms to find associations between quantitative imaging features and clinical outcome. This procedure is termed *Radiomics*, and it can be carried out in a variety of different ways. The most common approach is to use pre-defined, general purpose quantitative imaging features that describe intensity distribution, tumor shape, and heterogeneity. More modern technique, such as deep learning, allows the computer to learn problem-specific imaging features leading to more robust models.[49] Independently from the approach chosen, radiomics features can be combined with clinical and pathological data (possibly also extracted in the same fashion, i.e. pathomics) to predict clinical outcome, such as response to therapy [50]. Across most radiomics studies, it is noted how non-visual information relating to tumor heterogeneity is an important biomarker for response prediction.[51, 52] So far, radiomics has been applied to many tumor types (liver, lung, head and neck, brain) using varying modalities (CT, MRI, 18F-FDG PET/CT) with promising results.[50, 53, 54] Many studies have evaluated MR-based radiomics signatures[55-61], with a main focus on response evaluation. It is important to consider the technical challenges when applying radiomics on MR images: problems with standardization, normalization and regularization of images may hamper the generalizability of radiomics models.[62] Despite these difficulties, so far promising results

have been found in response prediction[56-59, 63] and response evaluation.[55, 60, 61, 64] Cui et al., for example, showed a favorable prognostic performance to predict pCR with radiomics on pre-CRT MR-images (AUC of 0.94-0.97)[56], but other studies reported lower accuracies (AUCs of 0.69-0.79)[57-59, 63] (Table 1). Van Griethuysen et al. showed that radiomics on pre-CRT MR images could predict response to therapy on image segmentation with comparable diagnostic performance as expert radiologists, regardless of their experience on image segmentation.[63] On post-CRT MR T2W-images, Horvat et al. showed that radiomics had a better classification performance compared to the combination of DWI and T2W-MRI to identify pCR, with a significant higher specificity and PPV (91% vs. 56% ; 72% vs 30%).[61] However, sensitivity and NPV were not significantly different. Another study concluded that radiomics could be used as an additional tool for clinical decision making on post-CRT imaging.[64] Until now, only single center studies using a heterogeneous methodology have been performed. Additionally, external validation of findings in radiomics research is often lacking, which is an important prerequisite to eventually apply developed predictive radiomics models in clinical practice. Currently, initiatives are being taken to deal with standardization of radiomics analyses and start up large datasets in order to facilitate external validation by international collaborations.

DISCUSSION

During the past decades advances in MR imaging technology and in image analysis and post processing methods have opened new windows of opportunity for research that will foster further personalization of treatment. Patients with smaller tumors will undergo neoadjuvant treatment with the aim to achieve a complete response and to offer organ preserving treatment. Main clinical questions concern the ability of a noninvasive imaging tool to accurately select before the onset of treatment those patients who are likely to achieve a (near) complete response. It is expected that functional parametric MRI will perform superior to conventional MRI because the combination of morphological and functional data provides a comprehensive information on the tumor. More advanced metrics derived from DWI perfusion and kurtosis imaging as well as Artificial intelligence modeling are promising but currently only the subject of research.

REFERENCES

1. Hermanek P, Wiebelt H, Staimmer D, Riedl S. Prognostic factors of rectum carcinoma--experience of the German Multicentre Study SGCRC. German Study Group Colo-Rectal Carcinoma. *Tumori*. May-Jun 1995;81(3 Suppl):60-4.
2. Maas M, Nelemans PJ, Valentini V, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol*. Sep 2010;11(9):835-44.
3. Joye I, Debucquoy A, Deroose CM, et al. Quantitative imaging outperforms molecular markers when predicting response to chemoradiotherapy for rectal cancer. *Radiother Oncol*. Jul 2017;124(1):104-109.
4. Beets-Tan RGH, Lambregts DMJ, Maas M, et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *Eur Radiol*. Apr 2018;28(4):1465-1475.
5. Martens MH, Van Heeswijk MM, Van Den Broek JJ, et al. Prospective, Multicenter Validation Study of Magnetic Resonance Volumetry for Response Assessment After Preoperative Chemoradiation in Rectal Cancer: Can the Results in the Literature be Reproduced? *International Journal of Radiation Oncology*Biology*Physics*. 2015;93(5):1005-1014.
6. Curvo-Semedo L, Lambregts DMJ, Maas M, et al. Rectal Cancer: Assessment of Complete Response to Preoperative Combined Radiation Therapy with Chemotherapy—Conventional MR Volumetry versus Diffusion-weighted MR Imaging. 2011;260(3):734-743.
7. Lambregts DM, Rao SX, Sassen S, et al. MRI and Diffusion-weighted MRI Volumetry for Identification of Complete Tumor Responders After Preoperative Chemoradiotherapy in Patients With Rectal Cancer: A Bi-institutional Validation Study. *Ann Surg*. Dec 2015;262(6):1034-9.
8. Intven M, Reerink O, Philippens ME. Diffusion-weighted MRI in locally advanced rectal cancer : pathological response prediction after neo-adjuvant radiochemotherapy. *Strahlenther Onkol*. Feb 2013;189(2):117-22.
9. Sun Y-S, Zhang X-P, Tang L, et al. Locally Advanced Rectal Carcinoma Treated with Preoperative Chemotherapy and Radiation Therapy: Preliminary Analysis of Diffusion-weighted MR Imaging for Early Detection of Tumor Histopathologic Downstaging. *Radiology*. 2010;254(1):170-178.
10. Jung SH, Heo SH, Kim JW, et al. Predicting response to neoadjuvant chemoradiation therapy in locally advanced rectal cancer: Diffusion-weighted 3 tesla MR imaging. *J Magn Reson Imaging*. 2012;35(1):110-116.
11. Joye I, Deroose CM, Vandecaveye V, Haustermans K. The role of diffusion-weighted MRI and 18F-FDG PET/CT in the prediction of pathologic complete response after radiochemotherapy for rectal cancer: A systematic review. *Radiother Oncol*. 2014;113(2):158-165.
12. Herneth AM, Guccione S, Bednarski M. Apparent diffusion coefficient: a quantitative parameter for in vivo tumor characterization. *Eur J Radiol*. Mar 2003;45(3):208-13.
13. Lambregts DMJ, Boellaard TN, Beets-Tan RGH. Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging: a pictorial review. *Insights into imaging*. Feb 13 2019;10(1):15.

14. Lu W, Jing H, Ju-Mei Z, et al. Intravoxel incoherent motion diffusion-weighted imaging for discriminating the pathological response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Sci Rep*. Aug 17 2017;7(1):8496.
15. Bakke KM, Hole KH, Dueland S, et al. Diffusion-weighted magnetic resonance imaging of rectal cancer: tumour volume and perfusion fraction predict chemoradiotherapy response and survival. *Acta Oncol*. 2017;56(6):813-818.
16. Hu F, Tang W, Sun Y, et al. The value of diffusion kurtosis imaging in assessing pathological complete response to neoadjuvant chemoradiation therapy in rectal cancer: a comparison with conventional diffusion-weighted imaging. *Oncotarget*. Sep 26 2017;8(43):75597-75606.
17. Yu J, Xu Q, Song JC, et al. The value of diffusion kurtosis magnetic resonance imaging for assessing treatment response of neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Eur Radiol*. May 2017;27(5):1848-1857.
18. De Cecco CN, Ganeshan B, Ciolina M, et al. Texture analysis as imaging biomarker of tumoral response to neoadjuvant chemoradiotherapy in rectal cancer patients studied with 3-T magnetic resonance. *Invest Radiol*. Apr 2015;50(4):239-45.
19. De Cecco CN, Ciolina M, Caruso D, et al. Performance of diffusion-weighted imaging, perfusion imaging, and texture analysis in predicting tumoral response to neoadjuvant chemoradiotherapy in rectal cancer patients studied with 3T MR: initial experience. *Abdominal radiology (New York)*. Sep 2016;41(9):1728-35.
20. Schurink NW, Lambregts DMJ, Beets-Tan RGH. Diffusion-weighted imaging in rectal cancer: current applications and future perspectives. *Br J Radiol*. Apr 2019;92(1096):20180655.
21. Tong T, Sun Y, Gollub MJ, et al. Dynamic contrast-enhanced MRI: Use in predicting pathological complete response to neoadjuvant chemoradiation in locally advanced rectal cancer. *J Magn Reson Imaging*. Sep 2015;42(3):673-80.
22. Intven M, Reerink O, Philippens ME. Dynamic contrast enhanced MR imaging for rectal cancer response assessment after neo-adjuvant chemoradiation. *J Magn Reson Imaging*. Jun 2015;41(6):1646-53.
23. Martens MH, Subhani S, Heijnen LA, et al. Can perfusion MRI predict response to preoperative treatment in rectal cancer? *Radiother Oncol*. Feb 2015;114(2):218-23.
24. Dijkhoff RAP, Maas M, Martens MH, et al. Correlation between quantitative and semiquantitative parameters in DCE-MRI with a blood pool agent in rectal cancer: can semiquantitative parameters be used as a surrogate for quantitative parameters? *Abdominal radiology (New York)*. May 2017;42(5):1342-1349.
25. Ciolina M, Caruso D, De Santis D, et al. Dynamic contrast-enhanced magnetic resonance imaging in locally advanced rectal cancer: role of perfusion parameters in the assessment of response to treatment. *Radiol Med*. May 2019;124(5):331-338.
26. Dijkhoff RAP, Beets-Tan RGH, Lambregts DMJ, Beets GL, Maas M. Value of DCE-MRI for staging and response evaluation in rectal cancer: A systematic review. *Eur J Radiol*. Oct 2017;95:155-168.
27. Kulkarni T, Gollins S, Maw A, Hobson P, Byrne R, Widdowson D. Magnetic resonance imaging in rectal cancer downstaged using neoadjuvant chemoradiation: accuracy of prediction of tumour stage and circumferential resection margin status. *Colorectal Dis*. Jun 2008;10(5):479-89.

28. Dresen RC, Beets GL, Rutten HJ, et al. Locally advanced rectal cancer: MR imaging for restaging after neoadjuvant radiation therapy with concomitant chemotherapy. Part I. Are we able to predict tumor confined to the rectal wall? *Radiology*. Jul 2009;252(1):71-80.
29. Allen SD, Padhani AR, Dzik-Jurasz AS, Glynne-Jones R. Rectal carcinoma: MRI with histologic correlation before and after chemoradiation therapy. *AJR Am J Roentgenol*. Feb 2007;188(2):442-51.
30. Ha HI, Kim AY, Yu CS, Park SH, Ha HK. Locally advanced rectal cancer: diffusion-weighted MR tumour volumetry and the apparent diffusion coefficient for evaluating complete remission after preoperative chemoradiation therapy. *Eur Radiol*. Dec 2013;23(12):3345-53.
31. Lambregts DMJ, Delli Pizzi A, Lahaye MJ, et al. A Pattern-Based Approach Combining Tumor Morphology on MRI With Distinct Signal Patterns on Diffusion-Weighted Imaging to Assess Response of Rectal Tumors After Chemoradiotherapy. *Dis Colon Rectum*. Mar 2018;61(3):328-337.
32. Lambregts DMJ, Maas M, Boellaard TN, et al. Long-term imaging characteristics of clinical complete responders during watch-and-wait for rectal cancer—an evaluation of over 1500 MRIs. *Eur Radiol*. 2020 Jan;30(1):272-280.
33. Mandard AM, Dalibard F, Mandard JC, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer*. Jun 1 1994;73(11):2680-6.
34. Patel UB, Taylor F, Blomqvist L, et al. Magnetic resonance imaging-detected tumor response for locally advanced rectal cancer predicts survival outcomes: MERCURY experience. *J Clin Oncol*. Oct 1 2011;29(28):3753-60.
35. Siddiqui MR, Gormly KL, Bhoday J, et al. Interobserver agreement of radiologists assessing the response of rectal cancers to preoperative chemoradiation using the MRI tumour regression grading (mrTRG). *Clin Radiol*. Sep 2016;71(9):854-62.
36. Battersby NJ, Dattani M, Rao S, et al. A rectal cancer feasibility study with an embedded phase III trial design assessing magnetic resonance tumour regression grade (mrTRG) as a novel biomarker to stratify management by good and poor response to chemoradiotherapy (TRIGGER): study protocol for a randomised controlled trial. *Trials*. Aug 29 2017;18(1):394.
37. Monguzzi L, Ippolito D, Bernasconi DP, Trattenero C, Galimberti S, Sironi S. Locally advanced rectal cancer: value of ADC mapping in prediction of tumor response to radiochemotherapy. *Eur J Radiol*. Feb 2013;82(2):234-40.
38. Ippolito D, Monguzzi L, Guerra L, et al. Response to neoadjuvant therapy in locally advanced rectal cancer: assessment with diffusion-weighted MR imaging and 18FDG PET/CT. *Abdom Imaging*. Dec 2012;37(6):1032-40.
39. Nougaret S, Vargas HA, Lakhman Y, et al. Intravoxel Incoherent Motion–derived Histogram Metrics for Assessment of Response after Combined Chemotherapy and Radiation Therapy in Rectal Cancer: Initial Experience and Comparison between Single-Section and Volumetric Analyses. *Radiology*. 2016;280(2):446-454.
40. Martens MH, Lambregts DMJ, Papanikolaou N, et al. Magnetization transfer imaging to assess tumour response after chemoradiotherapy in rectal cancer. *Eur Radiol*. 2016;26(2):390-397.
41. Martens MH, Lambregts DM, Papanikolaou N, et al. Magnetization transfer ratio: a potential biomarker for the assessment of postradiation fibrosis in patients with rectal cancer. *Invest Radiol*. Jan 2014;49(1):29-34.

42. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol*. Nov 2015;22(12):3873-80.
43. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
44. van der Sande ME, Beets GL, Hupkens BJ, et al. Response assessment after (chemo)radiotherapy for rectal cancer: Why are we missing complete responses with MRI and endoscopy? *Eur J Surg Oncol*. 2019 Jun;45(6):1011-1017.
45. Nahas SC, Rizkallah Nahas CS, Sparapan Marques CF, et al. Pathologic Complete Response in Rectal Cancer: Can We Detect It? Lessons Learned From a Proposed Randomized Trial of Watch-and-Wait Treatment of Rectal Cancer. *Dis Colon Rectum*. Apr 2016;59(4):255-63.
46. Smith FM, Wiland H, Mace A, Pai RK, Kalady MF. Clinical criteria underestimate complete pathological response in rectal cancer treated with neoadjuvant chemoradiotherapy. *Dis Colon Rectum*. Mar 2014;57(3):311-5.
47. Perez RO, Habr-Gama A, Pereira GV, et al. Role of biopsies in patients with residual rectal cancer following neoadjuvant chemoradiation after downsizing: can they rule out persisting cancer? *Colorectal Dis*. 2012;14(6):714-720.
48. Beets GL, Figueiredo NF, Beets-Tan RG. Management of Rectal Cancer Without Radical Resection. *Annu Rev Med*. Jan 14 2017;68:169-182. doi:10.1146/annurev-med-062915-021419
49. Trebeschi S, Van Griethuysen JJM, Lambregts DMJ, et al. Deep Learning for Fully-Automated Localization and Segmentation of Rectal Cancer on Multiparametric MR. *Sci Rep*. 2017 Jul 13;7(1):5301
50. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature communications*. Jun 3 2014;5:4006.
51. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. Feb 2016;278(2):563-77.
52. Kiessling F. The changing face of cancer diagnosis: From computational image analysis to systems biology. *Eur Radiol*. Aug 2018;28(8):3160-3164.
53. Bakr S, Echegaray S, Shah R, et al. Noninvasive radiomics signature based on quantitative analysis of computed tomography images as a surrogate for microvascular invasion in hepatocellular carcinoma: a pilot study. *Journal of medical imaging (Bellingham, Wash)*. Oct 2017;4(4):041303.
54. Zhou M, Scott J, Chaudhury B, et al. Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches. *AJNR Am J Neuroradiol*. Feb 2018;39(2):208-216.
55. Liu Z, Zhang X-Y, Shi Y-J, et al. Radiomics Analysis for Evaluation of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Clin Cancer Res*. 2017;23(23):7253-7262.
56. Cui Y, Yang X, Shi Z, et al. Radiomics analysis of multiparametric MRI for prediction of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Eur Radiol*. 2019;29(3):1211-1220.
57. Cusumano D, Dinapoli N, Boldrini L, et al. Fractal-based radiomic approach to predict complete pathological response after chemo-radiotherapy in rectal cancer. *Radiol Med*. 2018 Apr;123(4):286-295

58. Dinapoli N, Barbaro B, Gatta R, et al. Magnetic Resonance, Vendor-independent, Intensity Histogram Analysis Predicting Pathologic Complete Response After Radiochemotherapy of Rectal Cancer. *International Journal of Radiation Oncology*Biology*Physics*. 2018;102(4):765-774.
59. Meng Y, Zhang C, Zou S, et al. MRI texture analysis in predicting treatment response to neoadjuvant chemoradiotherapy in rectal cancer. *Oncotarget*. Feb 23 2018;9(15):11999-12008.
60. Bulens P, Couwenberg A, Haustermans K, et al. Development and validation of an MRI-based model to predict response to chemoradiotherapy for rectal cancer. *Radiother Oncol*. Mar 2018;126(3):437-442.
61. Horvat N, Veeraraghavan H, Khan M, et al. MR Imaging of Rectal Cancer: Radiomics Analysis to Assess Treatment Response after Neoadjuvant Therapy. *Radiology*. 2018;287(3):833-843.
62. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging*. Jan 2004;22(1):81-91.
63. van Griethuysen JJM, Lambregts DMJ, Trebeschi S, et al. Radiomics performs comparable to morphologic assessment by expert radiologists for prediction of response to neoadjuvant chemoradiotherapy on baseline staging MRI in rectal cancer. *Abdominal radiology (New York)*. 2020 Mar;45(3):632-643.
64. Bulens P, Couwenberg A, Intven M, et al. Predicting the tumor response to chemoradiotherapy for rectal cancer: Model development and external validation using MRI radiomics. *Radiother Oncol*. 2020 Jan;142:246-252.

CHAPTER 8

The use of deep learning on endoscopic images to assess the response of rectal cancer after chemoradiation

Hester E. Haak*, Xinpei Gao*, Monique Maas, Selam Waktola, Sean Benson, Regina G.H. Beets-Tan, Geerard L. Beets, Monique van Leerdam, Jarno Melenhorst

*** Contributed equally to this paper**

Surg Endosc. 2021 Oct 12.

ABSTRACT

Background

Accurate response evaluation is necessary to select complete responders (CRs) for a watch-and-wait approach. Deep learning may aid in this process, but so far has never been evaluated for this purpose. The aim was to evaluate the accuracy to assess response with deep learning methods based on endoscopic images in rectal cancer patients after neoadjuvant therapy.

Methods

Rectal cancer patients diagnosed between January 2012 and December 2015 and treated with neoadjuvant (chemo)radiotherapy were retrospectively selected from a single institute. All patients underwent flexible endoscopy for response evaluation. Diagnostic performance (accuracy, area under the receiver operator characteristics curve (AUC), positive- and negative predictive values, sensitivities and specificities) of different open accessible deep learning networks was calculated. Reference standard was histology after surgery, or long-term outcome (> 2 years of follow-up) in a watch-and-wait policy.

Results

226 patients were included for the study (117 (52%) were non-CRs; 109 (48%) were CRs). The accuracy, AUC, positive- and negative predictive values, sensitivity and specificity of the different models varied from 0.67-0.75, 0.76-0.83, 67%-74%, 70%-78%, 68%-79% and 66%-75% respectively. Overall, EfficientNet-B2 was the most successful model with the highest diagnostic performance.

Conclusions

This pilot study shows that deep learning has a modest accuracy (AUCs 0.76-0.83). This is not accurate enough for clinical decision making, and lower than what is generally reported by experienced endoscopists. Deep learning models can however be further improved and may become useful to assist endoscopists in evaluating the response. More well-designed prospective studies are required.

INTRODUCTION

Rectal cancer patients treated with neoadjuvant (chemo)radiotherapy (CRT) usually undergo reevaluation 6-10 weeks after the end of radiotherapy to evaluate therapy response. With an increasing interest in organ preservation, an additional goal of response evaluation is to identify a possible (near) complete response (CR). A combination of three modality assessment, digital rectal examination (DRE), endoscopy and MRI with diffusion-weighted imaging (DWI), has been shown to have the highest accuracy to identify a CR.[1] Many studies have addressed the value of MRI, while few studies have focused on endoscopy. Those which did evaluate the diagnostic value of endoscopy showed that it outperformed MRI in assessing the response.[1, 2] The majority of patients with a luminal CR (>70%) can be identified with endoscopy and a flat white scar is the most predictive feature to identify a CR.[2] However, 26 – 36% of the patients show other subtle morphological abnormalities, such as small or large flat ulcers, irregular tissue or residual adenomas, which are more difficult to interpret, leading to a considerable risk of missing residual disease or CRs.[1-3] New endoscopic techniques with computer aided diagnosis (CAD) using advanced imaging as narrow band imaging (NBI) or magnifying chromoendoscopy are designed to aid endoscopists in evaluating the histology of mucosal lesions, for example, by predicting submucosal invasions in advanced adenoma.[4] However, these techniques have not been studied in response assessment, and are limited due to the variability in diagnostic performance.[5] Other advances, in the field of artificial intelligence, in particular deep learning, may have potential to improve the endoscopic diagnostic accuracy.[5, 6] Deep learning neural networks uses many layers to automatically extract features. Automated methods such as deep learning are capable of analyzing large amounts of images at much faster rates than a human. It has already been shown to be effective in detecting small oesophageal cancer lesions[7, 8] or (benign) polyps in colon cancer[9-11] on endoscopy. The aim of this pilot study is to evaluate the feasibility and accuracy of deep learning methods based on endoscopic images and clinical variables for the response evaluation of rectal cancer patients treated with neoadjuvant therapy.

MATERIAL AND METHODS

Study design

The study cohort was retrospectively selected from a single institute database between January 2012 and December 2015. Informed consent was waived by the local institutional review board. Patients were included if they had 1) primary rectal cancer, 2) neoadjuvant long-course CRT or short course radiotherapy both followed by a waiting interval for downsizing, and 3) restaging endoscopic images available. Endoscopic restaging was routinely performed to assess the luminal response after neoadjuvant therapy. When residual disease was present at the response evaluation patients were referred for a total mesorectal excision (TME). When there was evidence of a clinical CR, patients were followed in a prospective watch-and-wait (W&W) study, approved by the local institutional review

board and registered on clinicaltrials.gov (NCT00939666 and NCT02278653). A clinical CR, as described in Maas et al.[12], consisted of no palpable tumor on DRE, white scar with no residual mass, ulcer or irregularity on endoscopy, and substantial downsizing with residual homogeneous fibrosis on T2-weigth imaging (T2W) without high signal on diffusion weighted imaging (DWI). Patients were excluded if they were: 1) lost to follow-up (FU), 2) refused surgery despite residual disease, or 3) maximum FU < 2 years when followed in a W&W program.

Endoscopy

All patients underwent flexible endoscopy (EPK-I video processor, Pentax Medical Netherlands, Uithoorn, the Netherlands) after neoadjuvant therapy to evaluate the luminal response. All patients received a rectal phosphate enema as a bowel preparation prior to endoscopy. Endoscopy was performed with standard white light imaging and the images (resolution of 768 x 576 pixels and 300 x 300 dpi) were digitally stored afterwards.

Predictive models

Model based on clinical variables

From the total of 226 patients, 70% (n = 158) were randomly allocated to a training/validation subset, 30% (n = 68) to a test subset, stratifying for CR and non-CR status. Three predictive models namely feedforward neural network (FFN), support vector machine (SVM) and logistic regression were built based on six clinical variables (age, sex, clinical T-stage, clinical N-stage, neoadjuvant treatment, and time between restaging endoscopy and surgery).

In order to reduce overfitting and improve the accuracy of the predictive models, the SelectKBest feature selection technique[13] was used to choose the best predicting clinical variables for the outcome (CR or non-CR). This technique scores all the features and then selects the optimal features according to the top highest scores. The top three selected clinical features (clinical N-stage, neoadjuvant treatment, and time between restaging endoscopy and surgery) were found to be optimum to train the models.[14] Performance of the clinical model was further assessed with the outcome measurements AUC, accuracy, precision, sensitivity and the F1 score. The F1-score is calculated as follows: F1Score=2*((precision*sensitivity)/(precision+sensitivity).

During training, 5-fold cross-validation was used on 70% of the training and validation set. The area under the curve (AUC) was calculated to assess model performance, where the loss function is minimized. During testing, bootstrapping calculated the model performance (AUC) of 500 randomly selected samples (with replacement) of the test subset. Mean AUC and the standard deviation of these 500 iterations were calculated to measure the model performance and the variability, respectively.

Deep learning based on endoscopy

Image preprocessing

A total of 731 endoscopic images were used which were split into training, validation and testing set, with the portion 7:3 resulting in 512 training/validation images and 219 test images. Since the number of available images were limited, 5-fold cross-validation was used to evaluate all the deep learning models. All endoscopic images belonging to the same patient were included in the same set. The median number of images per patient was 3 (range 2-7). The training set was used for the optimization of the weight of the neural network by the training process; the validation set was used to adjust the hyper-parameters (learning rate, number of epochs and size of mini batches) and the test set was independent from the training procedure, to test the final result of the neural network. The neural networks are trained using an optimization process that requires a loss function to calculate the error in the model. During training, if the prediction matches with the actual results the values of the loss function will be lower. The results of the independent test set will be presented. Figure 1 shows examples of the endoscopic images of CRs and non-CRs. The endoscopic images were first preprocessed to focus on the important features of the image (Figure 2). Preprocessing consisted of cutting the black margin of the images followed by cropping out the central region of the image. The images were also resized based on model image input size, rescaled and normalized. We applied a data augmentation procedure to increase the number of images used for the training set and to avoid overfitting.[15] The images for training were expanded by rotation, flipping, shearing and zooming of the original images, resulting in 4 additional images per patient. The number of CRs in the training/validation and test set were 76 and 33 patients, respectively.

Convolutional neural networks (CNN)

In order to develop an accurate deep learning algorithm large datasets are needed using a vast amount of data. An alternative way is to use transfer learning via ImageNet[16], which is an online accessible tool and re-uses networks which were trained by an enormous amount of natural images. Fine-tuning existing CNN models that have been pre-trained with a large set of labeled natural images are common to use as a transfer learning technique. The re-used network can be performed on medical data and show promising result. [17, 18] Specifically, models trained on the ImageNet dataset (~1.2 million training RGB images) could be suitable to apply transfer training learning while training endoscopic images where the available dataset is limited.[17] To improve the output of CNN, a classical method is used to increase the number of layers of CNN, which will also lead to the increase of time to compute and the difficulty to converge. As the number of layers increases, when the gradient is back propagated, multiple times of multiplication will make the gradient unstable, which is called a gradient explode or vanish problem. To improve this, novel structures are used. "ResNet" have shortcut connections between layers, and the output of previous layer will be added to the input of later layers.[19] ResNet's design is able to stabilize gradient and relieve the gradient explosion or vanishing problem. "DenseNet" has a similar concept with ResNet, and connects all the output of previous layers to the later layer by concatenating.[20]

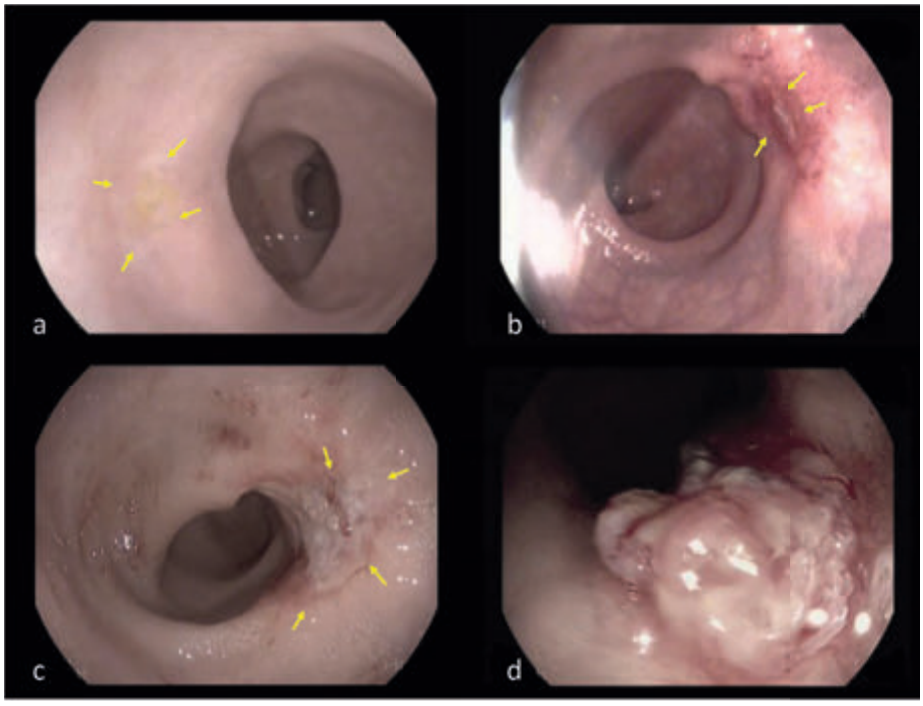


Figure 1 Example of evident and doubtful complete responders and non-complete responders. Evident complete response with a typical white scar (yellow arrows) (a), doubtful response with a small ulcer (yellow arrows) (b), doubtful response with a small-medium sized ulcer (yellow arrows) (c), and evident incomplete response with a tumor mass (d).

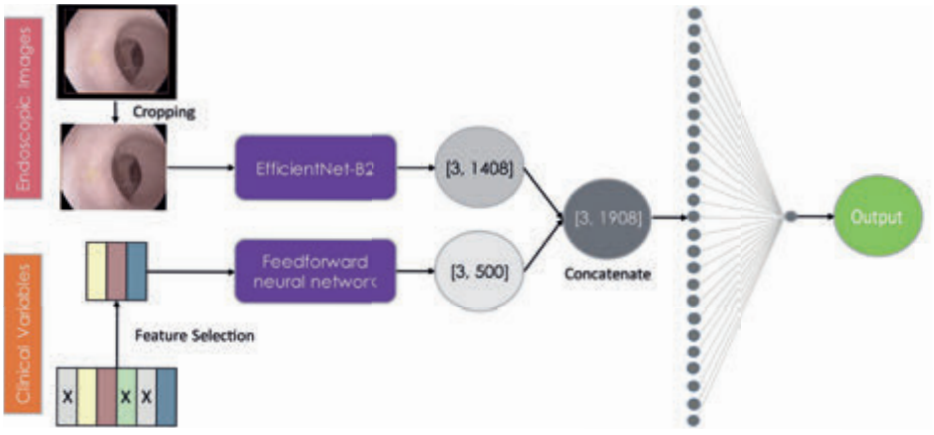


Figure 2 Overview of the combined model architecture. Preprocessing of endoscopic images for deep learning [1408] represents the last channels in EfficientNet-B2. [500] represents the number of neurons in feedforward neural network based on three selected clinical features.

DenseNet can also relieve the gradient explode or vanish problem and needs less computational time and memory compared to ResNet. Inception adds multiscale convolution modules in parallel, collects multiscale feature maps and concatenate, to increase the ability to learn feature representation.[21] Inception avoids the problem of adding too many layers. "InceptionResNet" combines the basic module of Inception and ResNet to achieve the advantages of both.[22] "Xception" is an improved version of Inception.[23] It replaces the convolution modules in Inception to depth wise separable convolutions, separating completely the relevance of channels. "MobileNet" separates convolution into depth wise and pointwise convolutions, compresses the network and also keeps the accuracy level[24] "EfficientNet" is one of the recent convolutional neural network architectures that achieve much better accuracy and efficiency as compared to the previous ConvNets. It uses a new scaling method that uniformly scales all dimensions of depth, width, and resolution to obtain a family of deep learning models.[25] In our study, we test and compared several CNNs including Xception, MobileNetV2, DenseNet121, ResNet50, InceptionV3, InceptionResNetV2 and EfficientNet-B2. They are mostly models with top results in natural object recognition in ImageNet Large Scale Visual Recognition Challenge(ILSVRC) competition.[26] The CNNs were trained with two 4GB K2 Nividia Graphics Processing Unit (GPU)s. The optimizer was Adam with the learning rate 1e-4. To find the best optimizer, we also tried SGD (stochastic gradient descent), RMSprop, and the learning rate was adjusted from 1e-3 to 1e-5. All the layers of the models were trained; and reducing the number of layers trained did not improve the performance. We also tried to change the training scheme, such as first training the bottom layer, then all the layers, which did not make a difference with current methods. The initialization weights of the models were the weight trained by ImageNet, the results became much worse with random initialized weights.

Combined model

Deep learning models in medical applications are increasingly combining contextual data from electronic health records and pixel data, because the clinical context is often crucial in diagnostic decisions.[27] Hence, in the present study the clinical features are combined with endoscopic imaging features to improve the performance of the deep learning models and provide more clinically relevant models. FFN was chosen to combine the selected clinical features (clinical N-stage, neoadjuvant treatment, and time between restaging endoscopy and surgery), in which it was the best performing model from the models constructed based on clinical variables. The late fusion technique[28] is used to train the combined models where the deep learning models extract features from the endoscopic images and the FFN part extracts features from the selected clinical variables. The combined model architecture is presented in Figure 2.

Reference standard

The outcome of the deep learning method was compared with the reference standard: non-CRs or CRs. In patients who were operated, the histopathological staging of the surgical resection specimen provided the reference standard, and in W&W patients follow-

up provided the reference standard. Non-CRs were either defined as patients who had residual luminal disease at histopathology (yT1-4) after resection, or W&W patients who developed a local regrowth (LR) during follow-up. CRs were defined as patients who had a pathological complete response (pCR) according histopathology (ypT0) or W&W patients with a sustained clinical CR after at least 2 years, as the vast majority of LRs occur within the first two years of FU [29, 30]. Because this study focused on the luminal response assessment, nodal stage was not included.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics version 22.0 (IBM Corporation 2013, Armonk, NY). For this pilot study no formal sample size calculation was made. Nominal data are presented as absolute frequencies and values and continuous data as median numbers with interquartile range (IQR). Baseline characteristics were compared between patients with and without a CR during FU. Differences were tested for significance with the χ^2 test for the comparison of proportions and the use of Mann Whitney-U-test for comparison of the medians. The diagnostic performance of the deep learning models was calculated by use of the following parameters: accuracy, area under the receiver operator characteristics curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) with 95% confidence intervals (CIs). The diagnostic performance of all the parameters was calculated according the binary outcome (CR or no CR), and CR was the positive outcome measure.

RESULTS

Demographics

238 patients were eligible for the study, of which 12 patients were excluded for the following reasons: W&W FU less than 2 years (n=2), refused surgery despite residual disease (n=5), lost to FU (n=2) and missing values (n=3). A total of n=226 patients were included in the analysis. Demographics of all patients are shown in Table 1. Median age was 65 (58-73) years and 153 (68%) of the patients were male. 206 (90%) of the patients received neoadjuvant CRT, the remaining 20 (10%) patients had short course radiotherapy with a prolonged waiting interval. In total, 117 (52%) of 226 patients had residual disease: 94 patients after immediate surgery and 23 patients in the W&W program who developed a regrowth (16 ypT1, 41 ypT2, 56 ypT3 and 4 ypT4). 109 (48%) of 226 patients were CRs: 19 with ypT0 after TME surgery, 85 W&W patients with a sustained ycT0 with a median FU of 53 months (26-69), and 5 W&W patients who underwent surgery for a suspected regrowth but did have a ypT0.

Variables	All (n=226)	Non-CR (n=117)	CR (n=109)	*P*
Age, median (IQR), yr	65 (58-73)	65 (58-74)	66 (59-73)	0.952
Sex, n (%)				
Male	153 (68)	78 (67)	75 (69)	0.731
Female	73 (32)	39 (33)	34 (31)	
Clinical T-stage, n(%)				
1-2	49 (22)	21 (18)	28 (26)	0.095
3	161 (71)	84 (72)	77 (70)	
4	16 (7)	12 (10)	4 (4)	
Clinical N-stage, n(%)				
0	54 (24)	26 (22)	28 (26)	0.038
1	64 (28)	26 (22)	38 (35)	
2	108 (48)	65 (56)	43 (39)	
Distance anal verge, n(%)				
≤ 5 cm	165 (73)	80 (68)	85 (78)	0.042
≥ 5 cm	61 (27)	37 (32)	24 (22)	
Neoadjuvant treatment, n(%)				
5x5Gy + prolonged waiting interval	20 (10)	16 (14)	4 (4)	<0.001
CRT	206 (90)	101 (86)	105 (96)	
Adjuvant chemotherapy, n(%)				
Yes	41 (18)	22 (19)	19 (17)	0.227
No	185 (82)	95 (81)	90 (83)	
Time between last radiotherapy and endoscopy, median (IQR), weeks	10 (8-15)	8 (8-12)	12 (9-18)	<0.001
Time between restaging endoscopy and surgery, median (IQR), weeks	5 (2-10)	4 (2-12)	6 (3-10)	0.359
Final treatment, n(%)				
W&W	113 (50)	23 (20)	90 (83)	<0.001
Immediate surgery	113 (50)	94 (80)	19 (17)	

Table 1 Patient characteristics of the total cohort and with and without a complete response during follow-up. CR=complete response; no-CR=no complete response, *P* = p-value; IQR=interquartile range; Gy=gray; CRT=chemoradiation; W&W=watch-and-wait

Performance of the models

In this section, we show the automatically generated results of models constructed based on clinical features, endoscopic images and combined (endoscopic images and clinical features) models of the same test set, which was independent from the training procedure to test the final result of the neural network, and compared the outcomes with the reference standard.

Machine learning models based on clinical features

Supplementary figure 1 summarizes the performance of the machine learning models built on all clinical features (sex, age, clinical T-stage, clinical N-stage, adjuvant chemotherapy,

and time between restaging endoscopy and surgery) and selected clinical features (clinical N-stage, adjuvant chemotherapy, and time between restaging endoscopy and surgery). The performance of models built on the three selected clinical features was higher than the model built on all clinical features. When considering the three selected clinical features, the FFN model performed slightly better (AUC of 0.73±0.05; accuracy of 0.70 ±0.04) than the SVM model (AUC 0.74±0.05; accuracy 0.68 ±0.04) and the logistic regression model (AUC 0.71±0.06; accuracy 0.64±0.04).

Deep learning models based on endoscopic images with and without clinical features

The performance of the different models using endoscopic images as an input was lower than the performance of the combined model in which imaging and clinical features were used. The AUCs for the different CNN models using endoscopic images only ranged from 0.71-0.79 and was best for EfficientNet-B2 with an AUC of 0.79 (95%CI 0.75-0.82) and a sensitivity of 0.74 (95%CI 0.70-0.78) and specificity of 0.70 (95%CI 0.66-0.74). All models based on endoscopic images only performed worse than the combined model. A detailed overview of the diagnostic performance of the endoscopic image models only are described in Supplementary table 1. The performance of the combined models are described in Table 2.

	Xception	MobileNet	DenseNet 121	ResNet50	InceptionV3	Inception ResNetV2	EfficientNet-B2
AUC	0.81	0.81	0.78	0.78	0.81	0.76	0.83
(95%CI)	(0.78-0.84)	(0.77-0.84)	(0.74-0.82)	(0.74-0.81)	(0.77-0.84)	(0.72-80)	(0.80-0.86)
Accuracy	0.75	0.70	0.69	0.67	0.72	0.69	0.75
(95%CI)	(0.71-0.79)	(0.66-0.74)	(0.65-0.73)	(0.63-0.71)	(0.68-0.76)	(0.65-0.73)	(0.72-0.79)
PPV	0.74	0.67	0.71	0.67	0.73	0.67	0.74
(95%CI)	(0.71-0.78)	(0.63-0.71)	(0.67-0.74)	(0.63-0.71)	(0.69-0.77)	(0.63-0.71)	(0.70-0.77)
NPV	0.78	0.70	0.71	0.72	0.74	0.71	0.77
(95%CI)	(0.74-0.80)	(0.66-0.74)	(0.67-0.75)	(0.68-0.75)	(0.70-0.77)	(0.67-0.75)	(0.74-0.80)
Sensitivity	0.79	0.76	0.68	0.73	0.71	0.68	0.77
(95%CI)	(0.75-0.82)	(0.73-0.80)	(0.64-0.72)	(0.70-0.77)	(0.67-0.75)	(0.64-0.72)	(0.73-0.80)
Specificity	0.73	0.73	0.72	0.66	0.72	0.71	0.75
(95%CI)	(0.69-0.77)	(0.70-0.77)	(0.68-0.76)	(0.62-0.70)	(0.69-0.76)	(0.67-0.75)	(0.72-79)

Table 2 Evaluation of the different convolutional neural network models including endoscopic images and clinical variables. CI= confidence interval; AUC=area under the ROC curve; PPV=positive predictive value; NPV=negative predictive value

The AUCs varied from 0.76-0.83 and was the highest in EfficientNet-B2 (0.83, 95%CI 0.80-0.86). Accuracy varied from 0.67-0.75, with EfficientNet-B2 and Xception having the highest accuracy (0.75, 95%CI 0.72-0.79; and 0.75, 95%CI 0.71-0.79 respectively). The PPV was the highest using EfficientNet-B2 (0.74, 95%CI 0.70-0.77) and Xception (0.74, 95%CI: 0.71-0.78) and varied from 67% to 74%. Xception had the highest NPV (0.78, 95% CI: 0.74-0.80)

and varied from 70% to 78%. Sensitivities varied from 68% to 79% and was the highest using Xception (0.79, 95%CI: 0.75-0.82). Specificities varied between 66% to 75%, and was the highest using EfficientNet-B2 (0.75, 95%CI 0.72-0.79). Figure 3 shows the diagnostic performance for EfficientNet-B2 and Supplementary figure 2 presents the loss value and accuracy of the training/validation datasets. Supplementary figure 3 gives an overview of the misclassified patients with EfficientNet-B2.

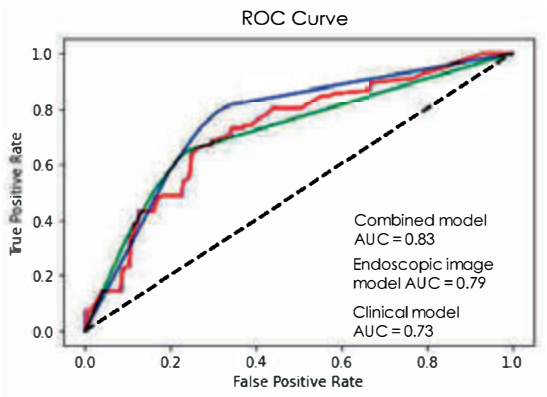


Figure 3 ROC-curve of EfficientNet-B2 for the endoscopic image model and combined model and ROC curve of feedforward neural network model for selected clinical variables. AUC=area under the ROC curve

DISCUSSION

The present study shows a modest accuracy of deep learning models based on both endoscopic images and clinical features to detect CRs on post-CRT endoscopy, and the combined model had a higher performance than models built on clinical features or endoscopic images only. The AUCs for the different CNN models ranged from 0.76 to 0.83. The CNN models detected 68% to 79% of the patients with a luminal CR. The diagnostic performance based on endoscopic images and clinical features varied between the different models but EfficientNet-B2 achieved the highest accuracy and AUC of 0.75 (95% CI 0.72-0.79) and 0.83 (95% CI: 0.80-0.86), respectively.

The AUCs in the present study are generally somewhat lower than the AUCs of 0.80 to 0.88 reported by experienced endoscopists by van der Sande et al. and Maas et al. [1, 2] The sensitivity of the AI models lies within the range of visual evaluation by endoscopists reported in the literature: 53% to 90%.[1, 2] Of course the sensitivity in the reported studies is highly dependent on the cut-off point, as illustrated by the study of Maas et al. where a

sensitivity of 53% was reported when using very strict selection criteria such as a white scar without any surface irregularities as a luminal CR. Currently, in many centers the selection criteria are less strict, leading to a higher sensitivity at the expense of a lower specificity to detect a CR. The specificity amongst the CNN models varied from 66% to 75%, generally somewhat lower than the 61% to 97% reported by experienced endoscopists.[1, 2]

Deep learning seems to be beneficial in other endoscopic areas, like in adenoma recognition where an algorithm correctly identifies diminutive (<5 mm) polyps in which a diagnose-and-leave strategy is accepted.[9, 11, 31, 32] Additionally, two randomized trials showed the efficacy of a real-time on-screen alert box in assisting endoscopists in polyp detection and evaluating the number of blind spots during procedures for quality measurements.[10, 33] In contrast to this, the current study showed a lower diagnostic value of the AI model than generally reported for expert endoscopists. Factors that may have contributed to this are the lack of high resolution images, and the input of only a limited number of 2D images per patient. High resolution images and real-time video assessment will likely lead to a higher performance. Moreover, the algorithm in the present study calculates the probability of a CR only on the basis of the few endoscopic images, whereas experienced endoscopists can also include the information of DRE and MR-imaging.[1] An additional limitation of the study is that some patients had a long interval between endoscopy and surgery, often because patients initially refused surgery or tried an alternative treatment. This may have caused a discrepancy between endoscopic images and histology of the resection specimen.

Clinical practice is shifting from providing a W&W approach in only typical cCR (a flat white scar without any surface irregularities at the first response assessment)[12, 34] to also selecting patients with a 'near CR' (ulcers, irregularity or adenoma) who can develop a flat white scar at a second reassessment after another interval.[35] The definition of this so called 'near CR' is unclear, and the subtle abnormalities are difficult to interpret with endoscopy only.[1-3] This is an area where we hope that deep learning methods can be of help. In addition, it could help endoscopists to guide focused biopsies in doubtful cases. In clinical practice the endoscopic assessment will never serve as a single-modality for decision making. Deep learning and AI can provide the endoscopic probability of a CR, and this information has to be added to the information of the other assessment methods. The addition of DRE and MRI-DWI to endoscopic evaluation had been shown to be highly valuable, with a particular value of MRI-DWI to detect in- and extra-luminal scattered tumor regions or nodal disease.[36] Clinical decision making is a complex process, that not only involves a probability estimate, but also patient and doctor preferences, and a number of other practical and ethical issues.[37, 38] Usually, physicians refer to non-imaging clinical data to interpret endoscopic imaging findings leading to higher diagnostic accuracy and more confident clinical decisions. EfficientNet-B2 demonstrated both higher accuracy and better efficiency over existing CNNs models and they also transfer well in multiple transfer learning datasets.[25] They also performed best in our endoscopic imaging and combined models, where EfficientNet-B2 had the highest performance. In order to develop an accurate deep learning model, a large amount of

data is required for collecting and labeling. However, when limited data is available, as in the current study, transfer learning has been shown to be a useful alternative, and there is evidence it even outperforms fully trained CNN models.[16, 17, 39, 40] To further explore the diagnostic performance of deep learning in response evaluation, additional studies are needed, such as multicenter cohorts evaluating a large amount of high resolution images or video material taken by different endoscopists. Possibly, adding other clinical input (e.g. DRE and MRI findings) can further improve the models.[41]

This retrospective pilot study shows that combining deep learning with clinical parameters to identify CR after neoadjuvant treatment for rectal cancer yields a diagnostic performance ranging from 0.76 to 0.83. The outcomes of CNN models varied widely, with EfficientNet-B2 being the most promising model. Compared to the literature, at present, an experienced endoscopist seems to be more accurate than deep learning. However, artificial intelligence may play a role in response evaluation when the performance of the models is further improved, and large prospective studies are required to explore this.

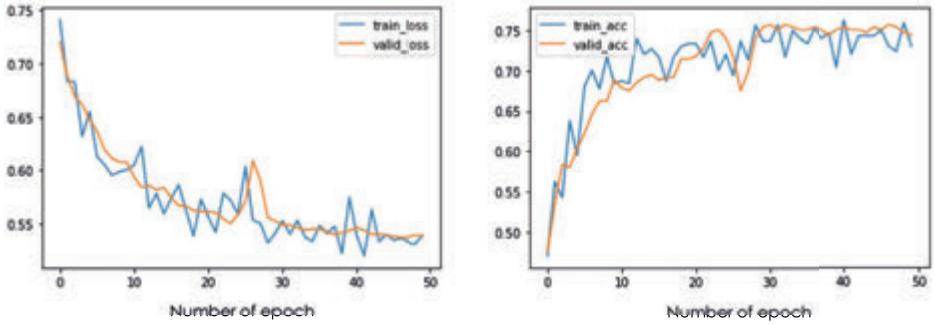
SUPPLEMENTARY MATERIAL

	All clinical features (sex, age, cT, cN, adjuvant chemotherapy, time between last radiotherapy and endoscopy)			Clinical features selected by SelectKBest (cN, adjuvant chemotherapy, time between last radiotherapy and endoscopy)		
	FFN	SVM	LogReg	FFN	SVM	LogReg
AUC (SD)	0.72 (0.05)	0.72 (0.05)	0.70 (0.06)	0.73 (0.05)	0.74 (0.05)	0.71 (0.06)
Accuracy (SD)	0.66 (0.05)	0.63 (0.04)	0.63 (0.03)	0.70 (0.04)	0.68 (0.04)	0.64 (0.04)
Sensitivity (SD)	0.66 (0.05)	0.63 (0.04)	0.63 (0.03)	0.70 (0.04)	0.68 (0.04)	0.64 (0.04)
Precision (SD)	0.67 (0.05)	0.65 (0.05)	0.65 (0.04)	0.71 (0.05)	0.69 (0.04)	0.66 (0.04)
F1-score	0.66	0.62	0.63	0.70	0.68	0.64

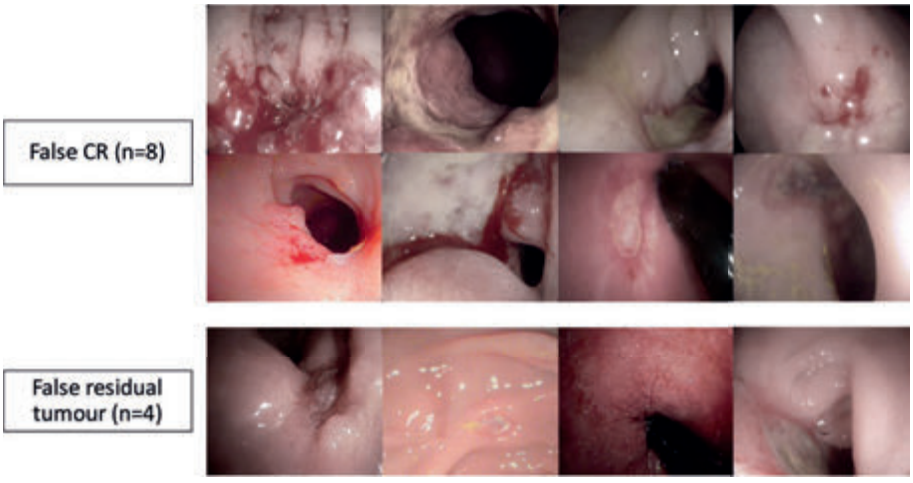
Supplementary figure 1 Clinical feature selection using SelectKBest with diagnostic performance of all the clinical features and the selected clinical features only. AUC=area under the ROC curve; SD=standard deviation; FFN=feedforward neural network; SVM=support vector machine; LogReg=logistic regression.

	Xception	MobileNet	DenseNet 121	ResNet50	Inception V3	Inception ResNetV2	EfficientNet-B2
AUC	0.76	0.73	0.75	0.71	0.74	0.72	0.79
(95%CI)	(0.73-0.80)	(0.69-0.76)	(0.71-0.78)	(0.67-0.75)	(0.70-0.77)	(0.68-0.76)	(0.75-0.82)
Accuracy	0.66	0.60	0.62	0.63	0.60	0.62	0.66
(95%CI)	(0.62-0.70)	(0.56-0.65)	(0.58-0.67)	(0.59-0.67)	(0.56-0.64)	(0.58-0.66)	(0.62-0.70)
PPV	0.69	0.65	0.65	0.54	0.63	0.66	0.63
(95%CI)	(0.64-0.72)	(0.60-0.69)	(0.61-0.70)	(0.50-0.58)	(0.59-0.67)	(0.62-0.70)	0.59-0.67)
NPV	0.70	0.64	0.66	0.69	0.65	0.62	0.68
(95%CI)	(0.66-0.73)	(0.60-0.68)	(0.62-0.70)	(0.65-0.73)	(0.61-0.69)	(0.58-0.66)	(0.64-0.72)
Sensitivity	0.64	0.63	0.67	0.68	0.65	0.64	0.74
(95%CI)	(0.60-0.68)	(0.59-0.67)	(0.63-0.71)	(0.64-0.72)	(0.61-0.69)	(0.59-0.68)	(0.70-0.78)
Specificity	0.70	0.52	0.62	0.65	0.65	0.66	0.70
(95%CI)	(0.66-0.74)	(0.47-0.56)	(0.58-0.67)	(0.61-0.70)	(0.61-0.69)	(0.62-0.70)	(0.66-0.74)

Supplementary table 1 Evaluation of the different convolutional neural network models including endoscopic images only. CI= confidence interval; AUC=area under the ROC curve; PPV=positive predictive value; NPV=negative predictive value



Supplementary figure 2 (a) the loss value and (b) accuracy of the training/validation dataset based on EfficientNet-B2 combined model.



Supplementary figure 3 Misclassified images of EfficientNet-B2 for the combined model.

REFERENCES

1. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol*. Nov 2015;22(12):3873-80. doi:10.1245/s10434-015-4687-9
2. van der Sande ME, Maas M, Melenhorst J, Breukink SO, van Leerdam ME, Beets GL. Predictive Value of Endoscopic Features for a Complete Response After Chemoradiotherapy for Rectal Cancer. *Ann Surg*. Dec 9 2019; doi: 10.1097/SLA.0000000000003718. Online ahead of print.
3. Nahas SC, Rizkallah Nahas CS, Sparapan Marques CF, et al. Pathologic Complete Response in Rectal Cancer: Can We Detect It? Lessons Learned From a Proposed Randomized Trial of Watch-and-Wait Treatment of Rectal Cancer. *Dis Colon Rectum*. Apr 2016;59(4):255-63.
4. Ferlitsch M, Moss A, Hassan C, et al. Colorectal polypectomy and endoscopic mucosal resection (EMR): European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy*. Mar 2017;49(3):270-297.
5. van der Sommen F, Curvers WL, Nagengast WB. Novel Developments in Endoscopic Mucosal Imaging. *Gastroenterology*. May 2018;154(7):1876-1886.
6. Ahmad OF, Soares AS, Mazomenos E, et al. Artificial intelligence and computer-aided diagnosis in colonoscopy: current evidence and future directions. *Lancet Gastroenterol Hepatol*. Jan 2019;4(1):71-80.
7. Hirasawa T, Aoyama K, Tanimoto T, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer*. Jul 2018;21(4):653-660.
8. de Groof AJ, Struyvenberg MR, van der Putten J, et al. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology*. Mar 2020;158(4):915-929.e4.
9. Byrne MF, Chapados N, Soudan F, et al. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut*. Jan 2019;68(1):94-100.
10. Wang P, Berzin TM, Glissen Brown JR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut*. Oct 2019;68(10):1813-1819.
11. Chen PJ, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology*. Feb 2018;154(3):568-575.
12. Maas M, Beets-Tan RG, Lambregts DM, et al. Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *J Clin Oncol*. Dec 10 2011;29(35):4633-40.
13. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12(85):2825-2830.
14. Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*. 2002/01/01 2002;46(1):389-422.
15. Abdollahi B, Tomita N, Hassanpour S. Data Augmentation in Training Deep Learning Models for Medical Image Analysis. In: Nanni L, Brahnam S, Brattin R, Ghidoni S, Jain LC, eds. *Deep Learners and Deep Learner Descriptors for Medical Applications*. Springer International Publishing; 2020:167-180.

16. Deng J, Dong W, Socher R, Li L, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. presented at: IEEE Conference on Computer Vision and Pattern Recognition; 20-25 June 2009 2009; Miami, FL, USA.
17. Tajbakhsh N, Shin JY, Gurudu SR, et al. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Trans Med Imaging*. May 2016;35(5)(5):1299-1312.
18. Shin HC, Roth HR, Gao M, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging*. May 2016;35(5):1285-98.
19. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016; Las Vegas, NV, USA.
20. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 09 November 2017 2017; Honolulu, HI, USA.
21. Szegedy C, Vanhoucke V, Loffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 12 December 2016 2016; Las Vegas, NV, USA.
22. Szegedy C, Ioffe S, vanhoucke V, Alemi A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI; 2017:
23. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017; Honolulu, HI, USA.
24. Howard AG, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *ArXiv*. 2017;
25. Tan M, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. presented at: Proceedings of the 36th International Conference on Machine Learning; 2019; Proceedings of Machine Learning Research. http://proceedings.mlr.press/v97/tan19a.html
26. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*. 2015/12/01 2015;115(3):211-252.
27. Huang S-C, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *npj Digital Medicine*. 2020/10/16 2020;3(1):136.
28. Ramachandram D, Taylor GW. Deep Multimodal Learning: A Survey on Recent Advances and Trends. *IEEE Signal Processing Magazine*. 2017;34(6):96-108.
29. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
30. Chadi SA, Malcomson L, Ensor J, et al. Factors affecting local regrowth after watch and wait for patients with a clinical complete response following chemoradiotherapy in rectal cancer (InterCoRe consortium): an individual participant data meta-analysis. *Lancet Gastroenterol Hepatol*. 2018 Dec;3(12):825-836.
31. Mori Y, Kudo SE, Misawa M, et al. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med*. Sep 18 2018;169(6):357-366.

32. Rex DK, Kahi C, O'Brien M, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc*. Mar 2011;73(3):419-22.
33. Wu L, Zhang J, Zhou W, et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut*. Dec 2019;68(12):2161-2169.
34. Martens MH, Maas M, Heijnen LA, et al. Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. *J Natl Cancer Inst*. Dec 2016;108(12).
35. Hupkens BJP, Maas M, Martens MH, et al. Organ Preservation in Rectal Cancer After Chemoradiation: Should We Extend the Observation Period in Patients with a Clinical Near-Complete Response? *Ann Surg Oncol*. Jan 2018;25(1):197-203.
36. van der Paardt MP, Zagers MB, Beets-Tan RG, Stoker J, Bipat S. Patients who undergo preoperative chemoradiotherapy for locally advanced rectal cancer restaged by using diagnostic MR imaging: a systematic review and meta-analysis. *Radiology*. Oct 2013;269(1):101-12.
37. Ragunath K. Artificial intelligence in gastrointestinal endoscopy: how intelligent can it get? *Lancet Oncol*. Dec 2019;20(12):1616-1617.
38. Le Berre C, Sandborn WJ, Aridhi S, et al. Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology*. Jan 2020;158(1):76-94.e2.
39. Brandao P, Zisimopoulos O, Mazomenos E, et al. Towards a Computed-Aided Diagnosis System in Colonoscopy: Automatic Polyp Segmentation Using Convolution Neural Networks. 2018;03(02):1840002.
40. Zhang R, Zheng Y, Wing Chung Mak T, Yu R, Wong SH, Lau JYW. Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. *IEEE Journal of Biomedical and Health Informatics* 05 December 2016 2017;21(1):41 - 47.
41. Bibault JE, Giraud P, Housset M, et al. Deep Learning and Radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. *Sci Rep*. Aug 22 2018;8(1):12611.

CHAPTER 9

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

In this thesis, currently pertinent challenges in organ preservation are evaluated regarding patient selection, follow-up and oncological and functional outcomes.

Patient selection

Selecting the right candidates for organ preservation is important. In case of obvious residual tumour at response evaluation a patient can proceed to surgery, but when patients have a good response elongating the interval between neoadjuvant treatment and response evaluation can make a difference. In these cases, when organ preservation is a specific aim, it makes sense to perform a second evaluation 6–12 weeks later, rather than proceed to total mesorectal excision (TME) surgery after the first restaging after chemoradiation (CRT) at 8-10 weeks in patients with a very good but not typical clinical complete response. Most watch-and-wait (W&W) centres perform response evaluation with digital rectal examination, magnetic resonance imaging (MRI) and endoscopy, which has been shown to be the most accurate approach to identify complete responders.[1, 2] Currently, all centres perform routine MRI restaging after neoadjuvant therapy, in agreement with the current guidelines.[3] Findings of our study *(chapter 2)* showed that with MRI the identification of poor responders was accurate, however, the identification of complete responders was more challenging. MRI has a tendency for overstaging and thereby misses complete responders because radiologists tend to err on the safe side to prevent that unrecognized residual disease in fibrosis will result in a regrowth.[4, 5] In addition, some complete responders have remaining mucosal irregularities, such as ulcers or adenomas on endoscopy or MRI abnormalities that look like residual tumour, but sometimes are not.[6, 7] Histology is generally considered as the gold standard, however, biopsies are of limited value given the high risk for false negatives and even false positives.[6, 8, 9] While endoscopy as a single-modality has the highest accuracy to identify luminal complete responders, MRI can preselect patients with an intermediate or good response for additional endoscopies. The following selection strategy might be adopted in clinical practice: in case of a poor response on MRI (with probably ycT3-4) patients can proceed to surgery without the need for endoscopy. In case of an intermediate or good response (ycT0-2) patients require an endoscopy to further evaluate the luminal response. When a small residual tumour (ycT1-2) is seen, patients could be directed for surgery or alternatively be treated with local excision or contact brachytherapy.[10-13] Also, a test of time with a second evaluation could be applied in patients with small residual disease or a near complete response to evaluate if a complete response will be achieved. When a complete response is seen patients are eligible to follow a W&W approach.

After neoadjuvant therapy, sterilization of lymph nodes is an important prerequisite for safe organ preservation. MRI aids in detecting residual lymph nodes metastases, however, the detection of lymph node metastases is challenging *(chapter 2)*.[4, 14] For nodal restaging and patient counseling knowledge about the prevalence of lymph node metastases after neoadjuvant therapy is crucial. Several studies reported outcomes on ypN-status per ypT-stage in different settings, but mainly in small single centre cohorts. *Chapter 3* provides important data on the prevalence of nodal disease after CRT and surgery, which is currently

difficult to obtain due to the increase in organ preserving strategies instead of TME surgery. The risk of lymph node metastases increases with increasing depth of residual tumour, and is per ypT-stage in the same range as per pT-stage in non-irradiated tumours. Although the risk of lymph node metastases is low in patients with ypT0, W&W studies report an even lower nodal regrowth rate.[1] This is probably due to a longer waiting interval between the end of neoadjuvant therapy and response evaluation which causes a further regression of nodal micrometastases. In addition, patients who have a good response after neoadjuvant therapy probably initially had a lower tumour volume at diagnosis which is associated with a lower risk of nodal metastases compared to patients with a large tumour volume who respond less.[15, 16] Another explanation for the low nodal regrowth rate in W&W patients is that nowadays MRI is routinely adopted and patients with obvious malignant lymph nodes or other high-risk tumour features such as extramural vascular invasion or MRF invasion are less likely to undergo W&W and instead are treated with surgery.

Clinically, the results of *chapter 2* and *chapter 3* can help guide response evaluation and treatment decision making. With this information W&W can be more accessible for both high and low volume centres who have less access to both resources, for example due to economical or logistical reasons. However, it should be kept in mind that when a (near) complete response is seen, patients should be referred to high volume W&W expert centres to ensure optimal treatment. In addition, patient counseling is key and several factors should be communicated when deciding for the best therapy in individual patients. First, the possibility of a complete response might influence treatment decision making, and can be estimated and discussed at primary diagnosis according to the tumour stage and extraluminal disease, taking into account the preference of the patient. After neoadjuvant therapy the possibility of a complete response can be estimated by response evaluation. Again, other factors such as patient preferences, risk for a permanent stoma and tumour location may influence treatment decisions. For example, when the tumour is located distally (<5 cm of the anal verge) there is a higher risk for an impaired functional outcome due to a low anastomosis or TME with a permanent stoma, and these patients tend to be more interested in organ preservation than patients with a proximal tumour.

Follow-up

When a patient is selected for W&W, frequent follow-up is performed to identify local regrowths early. There is no evidence regarding the efficacy of the currently employed follow-up schedules in W&W and therefore, schedules are highly variable. In The Netherlands patients are evaluated 3-monthly with MRI and endoscopy in the first year after inclusion for W&W and 6-monthly thereafter until completion of follow-up after five years. *Chapter 4* showed that increasing the frequency of endoscopy in the first two years and deintensifying the schedule after two years would make the follow-up schedule more efficient with regard to local regrowth detection, without the risk of missing late regrowths. It is expected that this new schedule will cause less burden for the patients and will be more cost-efficient. Follow-up schedules can be fine-tuned according to experience per centre: e.g. less experienced

centres can consider to evaluate more frequently if desired. A more intensive follow-up schedule is also advisable in patients with a higher risk to develop a regrowth, for example, in patients with a near complete response or a tumour remnant who undergo LE or additional contact brachytherapy. Recently, Dutch guidelines deintensified the standard follow-up after rectal resection from regular carcinoembryonic antigen (CEA) measurements and yearly computed tomography (CT) to regular CEA measurements and only performing CT once after 12 months or on indication.[17] The question arises whether a deintensified W&W follow-up schedule in combination with the new standard surveillance with CEA only will be accurate enough to detect local regrowths early and achieve similar oncological outcomes as in patients who are treated with a TME. Although the risk of nodal regrowth is very low, reducing the number of CTs might increase the risk for undetected nodal regrowths such as those along the inferior mesenteric vessels, which can be outside the field of view of the MRI. Multicentre prospective cohort studies who follow the new standard surveillance are needed to compare the deintensified W&W schedule with the current W&W schedule.

Oncological outcomes

W&W patients have excellent long-term outcomes with high overall survival, low risk of metastasis and low risk of locally unsalvageable disease.[1, 18] Due to aging of the population more patients 75 years of age or older will be diagnosed with rectal cancer. Especially for these patients the risk for morbidity, mortality and loss of independency after rectal surgery is high. W&W appears to be a safe alternative in older patients with a high pelvic control rate and few related rectal cancer deaths *(chapter 5)*. In addition, most patients avoided surgery and a permanent stoma and had reasonable anorectal and urinary function. It should however also be noted that due to improvements in frailty assessment, operative techniques and perioperative care, the risk of postoperative mortality in elderly patients has markedly improved.[19] Age by itself should not be a reason to withhold TME surgery if that is the best treatment in a given situation. In order to further evaluate the benefit of W&W specific studies on frail and very old patients are needed. A phase 3 randomized trial that is currently including patients evaluates organ preserving therapies in older and frail patients by assessing the added value of brachytherapy boost after neoadjuvant CRT. Endpoints of the study are clinical complete response at 6 months and oncological and functional outcomes, and results are expected in the upcoming years (NL69261.058.19).

Despite beneficial outcomes in W&W patients, it has been reported that patients who develop a local regrowth have a higher risk of metastases.[20, 21] One hypothesis is that true complete responders have more favorable tumour biology than patients with a regrowth who by definition actually did not have a complete response. Patients with a regrowth have inherently a more aggressive tumour with a higher risk of rapid metastases. Another hypothesis is that metastases arise from regrowths, the so-called “second hit theory”. Patients who develop regrowths are believed to harbor small residual clusters of radioresistant cancer cells in the bowel wall, not recognized by the response evaluation. This population of radioresistant cancer cells grows over time and may be biologically

more unfavourable than the primary tumour, with a higher chance to develop secondary metastases. A study that evaluates if the time to metastases in W&W patients is delayed compared to patients treated with CRT and surgery could provide evidence for either of these two hypotheses. A randomized controlled trial in patients with a complete response comparing the incidence of metastases of the whole group of patients who follow a W&W approach to a similar group undergoing TME would be the preferred study design. Although such a trial would provide definite evidence, inclusion would be difficult as patients are unwilling to take part in a study where the possibility for organ preservation is determined by chance. Another option is to perform a study comparing regrowths in patients in the W&W approach with patients treated with CRT and surgery who have residual disease at histopathology. However, due to several biases (immortal time bias and stage migration) such a comparison is methodologically challenging. As a next best option in *chapter 6* metastatic time patterns in W&W patients were compared with those in patients treated with CRT and surgery. The risk of metastases in W&W patients is lower than in operated patients, and this difference changes over time during the follow up. The difference is smaller in the later follow-up period compared to the early follow-up period. This suggests that distant metastases in W&W patients are generally detected later than in patients who are treated with CRT and surgery, and supports the "second hit theory" as a possible mechanism for at least a part of the metastases. Although unfortunately a definite answer could not be provided, it cannot be excluded that some of the metastases can arise from regrowths, and this potential small risk should be communicated during the shared decision making process when deciding for W&W or surgery. In addition, attempts should be made to minimize the risk for development of regrowths. More accurate selection with strict clinical and radiological criteria of patients with a true complete response will decrease the number of regrowths in a W&W program and will expose fewer patients to the risk of "second hit" metastases. The combination of digital rectal examination, endoscopy and MRI has the highest accuracy to detect complete responders with a local regrowth rate of 15-25%. Additional local excisions could prevent a number of regrowths but cause morbidity and the exact role needs to be determined. Potentially, advanced imaging techniques which are currently being developed might improve the selection process and consequently reduce the risk of regrowths and metastases.

Further improvements for successful organ preservation

Advanced imaging is expected to improve the selection of the right patient for the right treatment. *Chapter 7* gives an overview of current advanced MR imaging techniques and gives directions for further research in this area. In general, functional imaging outperforms conventional morphological imaging, because it provides comprehensive information of the tumour by capturing changes in tumour perfusion and microstructure before they become apparent on morphological imaging. More advanced imaging processing technologies such as radiomics and convolutional neural networks (CNNs) are also promising. These methods use computer algorithms to find associations with quantitative features (i.e. pre-defined general purpose features such as intensity distribution, tumour shape and

heterogeneity) and clinical outcome. A more modern technique is deep learning where a computer identifies problem-specific features leading to more robust models. Deep learning uses a CNN to extract features and link them together on a large scale to build predictive models. Unfortunately, *chapter 8* showed that the AUCs of experienced endoscopists are generally somewhat higher than the AUCs of the deep learning models. Factors that may have contributed to a lower performance of deep learning was that models were built on only a few low resolution images per patient with the addition of some clinical features. Endoscopists, however, have access to information on digital rectal examination and MR imaging which likely help with interpreting the response. In clinical practice the endoscopic assessment will never serve as a single-modality for decision making. A promising adjunct of deep learning might be to guide focused biopsies in patients were the response is unclear.

Another important area of research that can be further explored is response prediction. The ultimate goal of response prediction is that with the combination of multifactorial models including (advanced) imaging, clinical factors and blood-tissue biomarkers the best possible treatment can be found according to specific characteristics of a patient's tumour. This is called precision medicine were patients are stratified for treatment according to a molecular subtype and predicted response to therapy. For the past decade the number of tools and techniques to predict response and outcome by biomarkers increased significantly. Currently, when a patient is diagnosed with colorectal cancer the presence of specific genetic biomarkers such as microsatellite instability (MSI) / mismatch repair (MMR), RAS or BRAF mutation is often determined. The identification of these biomarkers can be used for treatment decisions on targeted therapy and immunotherapy. Rather than identifying single biomarkers, broader tests such as whole-genome sequencing can identify many more potentially actionable markers. In addition, several advanced imaging methods such as radiomics or CNNs are able to quantitatively evaluate features that can be used as predictive and prognostic biomarkers. Moreover, with radiogenomics quantitative or qualitative imaging features are linked to genomic profiles.[22, 23] The advantage of radiogenomics is that it can assess the biological profile and prediction of response in a non-invasive way of the whole tumour, without requiring a biopsy. The large bulk of evidence is in brain and lung tumours but some studies report outcomes of radiogenomics in rectal cancer.[24] Studies using CT or MRI based radiomic signatures were capable of predicting KRAS and BRAF mutations in patients with rectal cancer.[25, 26] Other studies using positron emission tomography CT with fluorodeoxyglucose 18F (F18 FDG PET/CT) imaging describe conflicting results on the prediction of KRAS mutational status.[27, 28] Overall, several steps need to be taken before such techniques (radiomics/radiogenomics or CNNs such as deep learning) can be adopted in routine clinical care. Further large-scale prospective studies are required. In addition, the majority of artificial intelligence studies (such as our deep learning study) have limitations regarding standardization, normalization, validation and regularization that hampers the generalizability. Therefore, open access datasets need to be provided that are broadly available for international centres to allow for standardization of different techniques and multicentre external validation. Nevertheless, the results as presented in

chapter 7 and 8 support to further pursue this line of research, while addressing the issues with regard to study design and methodology as mentioned above. For the realization of such a high level of personalized medicine, patients need to be treated in an institute of integrated diagnosis where clinicians from different specialties work together with computer scientists and biotechnicians. However, to reach this goal it is essential that centres have the availability of the right infrastructure, resources and expertise.

In addition, improvements in response prediction, systemic therapy and radiotherapy will also increase the chance for organ preservation. For example, patients who have a higher risk for a poor oncological outcome might be candidates for a more intense neoadjuvant strategy incorporating systemic treatment, the so called total neoadjuvant therapy (TNT). TNT schedules (chemo + CRT) increases the compliance and leads to more complete responses and organ preservation. A recent systematic review and meta-analysis reported an OR of 2.44 for the TNT group to achieve a pathological complete response compared to patients who received neoadjuvant CRT and adjuvant chemotherapy.[29] In addition, preliminary results of the OPRA trial reported a 3-year organ preservation rate of 43% in patients who received induction chemotherapy and CRT compared to 58% in patients who received CRT and consolidation chemotherapy.[30] Another approach is to increase the radiotherapy dose, for example with dose escalation[31, 32], external boosting[33] or endoluminal boosting[34-37]. Also, MR linac may increase the effectiveness of radiotherapy by daily response monitoring during neoadjuvant treatment which provides the opportunity to change treatment according to response, while avoiding unwanted dose on non-target organs.[38] Furthermore, due to the national implementation study of the W&W strategy more patients want to undergo CRT to achieve a complete response. Therefore, neoadjuvant treatment might be considered in early disease, because small tumours have a higher chance of complete response than locally advanced tumours.[15, 16] Also, local surgical therapy (e.g. TEM, TAMIS) in good or complete responders after CRT instead of direct TME might be an interesting alternative to achieve organ preservation. The TESAR trial (NCT02371304) and STARTREC trial (NCT02945566) are currently recruiting patients with early disease to evaluate these treatment options.

What the future holds

Ideally, the following scenario will become reality in 10 years' time. A patient who has a suspicion of a rectal tumour undergoes a colonoscopy. A luminal rectal tumour mass is seen at 5 cm of the anal verge, a biopsy is taken to establish the diagnosis and to provide a full molecular classification of the tumour through whole-genome sequencing. The patient undergoes (advanced) MR imaging for anatomical evaluation of the relation of the tumour to the layers of the bowel wall and the surrounding structures, and with a functional evaluation of (imaging) biomarkers that predict response to different therapies. All this information comes together at a multidisciplinary team meeting with the involved clinicians, physician assistants, biotechnicians and computer scientists. According to tumour type, tumour staging and biomarkers the likelihood for response is estimated. The different treatment options are listed, with some treatment options focusing more on the highest chance of

cure, other options focusing more on good long term quality of life, and again other options focusing on the least toxicity of the treatment. These options are then discussed with the patient taking into account a patient's age, lifestyle and expectations, and the risks and benefits of the different options. This often will take more than one visit, but the time invested in shared decision making will pay off in a more satisfied patient with the most effective treatment on the long term.

REFERENCES

1. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. Jun 23 2018;391(10139):2537-2545.
2. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol*. Nov 2015;22(12):3873-80.
3. Beets-Tan RGH, Lambregts DMJ, Maas M, et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *Eur Radiol*. Apr 2018;28(4):1465-1475.
4. van der Paardt MP, Zagers MB, Beets-Tan RG, Stoker J, Bipat S. Patients who undergo preoperative chemoradiotherapy for locally advanced rectal cancer restaged by using diagnostic MR imaging: a systematic review and meta-analysis. *Radiology*. Oct 2013;269(1):101-12.
5. Lambregts DMJ, Boellaard TN, Beets-Tan RGH. Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging: a pictorial review. *Insights into imaging*. Feb 13 2019;10(1):15.
6. van der Sande ME, Beets GL, Hupkens BJ, et al. Response assessment after (chemo)radiotherapy for rectal cancer: Why are we missing complete responses with MRI and endoscopy? *Eur J Surg Oncol*. Jun 2019;45(6):1011-1017.
7. Smith FM, Chang KH, Sheahan K, Hyland J, O'Connell PR, Winter DC. The surgical significance of residual mucosal abnormalities in rectal cancer following neoadjuvant chemoradiotherapy. *Br J Surg*. Jul 2012;99(7):993-1001.
8. Perez RO, Habr-Gama A, Pereira GV, et al. Role of biopsies in patients with residual rectal cancer following neoadjuvant chemoradiation after downsizing: can they rule out persisting cancer? *Colorectal Dis*. 2012;14(6):714-720.
9. Smith FM, Wiland H, Mace A, Pai RK, Kalady MF. Clinical criteria underestimate complete pathological response in rectal cancer treated with neoadjuvant chemoradiotherapy. *Dis Colon Rectum*. Mar 2014;57(3):311-5.
10. Rullier E, Vendrely V, Asselineau J, et al. Organ preservation with chemoradiotherapy plus local excision for rectal cancer: 5-year results of the GRECCAR 2 randomised trial. *Lancet Gastroenterol Hepatol*. May 2020;5(5):465-474.
11. Borschitz T, Wachtlin D, Mohler M, Schmidberger H, Junginger T. Neoadjuvant chemoradiation and local excision for T2-3 rectal cancer. *Ann Surg Oncol*. Mar 2008;15(3):712-20.
12. Sun Myint A, Smith FM, Gollins S, et al. Dose Escalation Using Contact X-ray Brachytherapy After External Beam Radiotherapy as Nonsurgical Treatment Option for Rectal Cancer: Outcomes From a Single-Center Experience. *Int J Radiat Oncol Biol Phys*. Mar 1 2018;100(3):565-573.
13. Sun Myint A, Smith FM, Gollins SW, et al. Dose escalation using contact X-ray brachytherapy (Papillon) for rectal cancer: does it improve the chance of organ preservation? *Br J Radiol*. Dec 2017;90(1080):20170175.

14. Bipat S, Glas AS, Slors FJ, Zwinderman AH, Bossuyt PM, Stoker J. Rectal cancer: local staging and assessment of lymph node involvement with endoluminal US, CT, and MR imaging--a meta-analysis. *Radiology*. Sep 2004;232(3):773-83.
15. Jankowski M, Pietrzak L, Rupinski M, et al. Watch-and-wait strategy in rectal cancer: Is there a tumour size limit? Results from two pooled prospective studies. *Radiother Oncol*. Jul 2021;160:229-235.
16. Gérard J-P, Chamorey E, Gourgou-Bourgade S, et al. Clinical complete response (cCR) after neoadjuvant chemoradiotherapy and conservative treatment in rectal cancer. Findings from the ACCORD 12/PRODIGE 2 randomized trial. *Radiother Oncol*. 2015/05/01/ 2015;115(2):246-252.
17. Richtlijnendatabase Federatie medisch specialisten: Follow-up na chirurgische resectie stadium I-III colon- en rectumcarcinoom 2021. https://richtlijnendatabase.nl/richtlijn/colorectaal_carcinoom_crc/follow-up_bij_crc/follow-up_na_chirurgische_resectie_stadium_i-iii_colon-_en_rectumcarcinoom.html
18. van der Sande ME, Figueiredo N, Beets GL. Management and Outcome of Local Regrowths in a Watch-and-wait Prospective Cohort for Complete Responses in Rectal Cancer. *Ann Surg*. Feb 7 2020; doi: 10.1097/SLA.0000000000003738. Online ahead of print.
19. Brouwer NPM, Heil TC, Olde Rikkert MGM, et al. The gap in postoperative outcome between older and younger patients with stage I-III colorectal cancer has been bridged; results from the Netherlands cancer registry. *Eur J Cancer*. Jul 2019;116:1-9.
20. Smith JJ, Strombom P, Chow OS, et al. Assessment of a Watch-and-Wait Strategy for Rectal Cancer in Patients With a Complete Response After Neoadjuvant Therapy. *JAMA oncology*. Jan 10 2019:e185896.
21. Socha J, Kępka L, Michalski W, Paciorek K, Bujko K. The risk of distant metastases in rectal cancer managed by a watch-and-wait strategy – A systematic review and meta-analysis. *Radiother Oncol*. 2020;144:1-6.
22. Horvat N, Bates DDB, Petkovska I. Novel imaging techniques of rectal cancer: what do radiomics and radiogenomics have to offer? A literature review. *Abdominal Radiology*. 2019/11/01 2019;44(11):3764-3774.
23. Bodalal Z, Trebeschi S, Beets-Tan R. Radiomics: a critical step towards integrated healthcare. *Insights into imaging*. 2018/12/01 2018;9(6):911-914. doi:10.1007/s13244-018-0669-3
24. Bodalal Z, Trebeschi S, Nguyen-Kim TDL, Schats W, Beets-Tan R. Radiogenomics: bridging imaging and genomics. *Abdominal Radiology*. 2019/06/01 2019;44(6):1960-1984.
25. Yang L, Dong D, Fang M, et al. Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? *Eur Radiol*. 2018/05/01 2018;28(5):2058-2067.
26. Shin YR, Kim KA, Im S, Hwang SS, Kim K. Prediction of KRAS Mutation in Rectal Cancer Using MRI. *Anticancer Res*. Sep 2016;36(9):4799-804.
27. Lovinfosse P, Koopmansch B, Lambert F, et al. 18F-FDG PET/CT imaging in rectal cancer: relationship with the RAS mutational status. *The British Journal of Radiology*. 2016;89(1063):20160212.
28. Chen SW, Lin CY, Ho CM, et al. Genetic Alterations in Colorectal Cancer Have Different Patterns on 18F-FDG PET/CT. *Clin Nucl Med*. Aug 2015;40(8):621-6.
29. Kasi A, Abbasi S, Handa S, et al. Total Neoadjuvant Therapy vs Standard Therapy in Locally Advanced Rectal Cancer: A Systematic Review and Meta-analysis. *JAMA Netw Open*. Dec 1 2020;3(12):e2030097.

30. Garcia-Aguilar J, Patil S, Kim JK, et al. Preliminary results of the organ preservation of rectal adenocarcinoma (OPRA) trial. *J Clin Oncol*. 2020;38(15_suppl):4008-4008.
31. Appelt AL, Ploen J, Vogelius IR, Bentzen SM, Jakobsen A. Radiation dose-response model for locally advanced rectal cancer after preoperative chemoradiation therapy. *Int J Radiat Oncol Biol Phys*. Jan 1 2013;85(1):74-80.
32. Habr-Gama A, São Julião GP, Vailati BB, et al. Organ Preservation in cT2N0 Rectal Cancer After Neoadjuvant Chemoradiation Therapy: The Impact of Radiation Therapy Dose-escalation and Consolidation Chemotherapy. *Ann Surg*. Jan 2019;269(1):102-107.
33. Couwenberg AM, Burbach JPM, Berbee M, et al. Efficacy of Dose-Escalated Chemoradiation on Complete Tumor Response in Patients with Locally Advanced Rectal Cancer (RECTAL-BOOST): A Phase 2 Randomized Controlled Trial. *Int J Radiat Oncol Biol Phys*. Nov 15 2020;108(4):1008-1018.
34. Ortholan C, Romestaing P, Chapet O, Gerard JP. Correlation in rectal cancer between clinical tumor response after neoadjuvant radiotherapy and sphincter or organ preservation: 10-year results of the Lyon R 96-02 randomized trial. *Int J Radiat Oncol Biol Phys*. Jun 1 2012;83(2):e165-71.
35. Appelt AL, Pløen J, Harling H, et al. High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. *The Lancet Oncology*. 2015/08/01/ 2015;16(8):919-927.
36. Rijkmans EC, Cats A, Nout RA, et al. Endorectal Brachytherapy Boost After External Beam Radiation Therapy in Elderly or Medically Inoperable Patients With Rectal Cancer: Primary Outcomes of the Phase 1 HERBERT Study. *Int J Radiat Oncol Biol Phys*. Jul 15 2017;98(4):908-917.
37. Gérard JP, Barbet N, Dejean C, et al. Contact X-ray brachytherapy for rectal cancer: Past, present, and future. *Cancer/Radiothérapie*. 2021 May 26;S1278-3218(21)00077-9.
38. Boldrini L, Intven M, Bassetti M, Valentini V, Gani C. MR-Guided Radiotherapy for Rectal Cancer: Current Perspective on Organ Preservation. Perspective. *Front Oncol*. 2021-March-30 2021;11(530).

APPENDICES

Summary

The aims of this thesis are: 1] to evaluate if magnetic resonance imaging (MRI) alone can accurately identify patients who have substantial residual disease after neoadjuvant chemoradiation (CRT) that requires immediate surgery, and if these findings are reproducible amongst radiologists with various expertise levels; 2] to evaluate the pooled prevalence of lymph nodes after CRT according to increasing depth of residual tumour in the rectal wall and to assess the impact of post-CRT lymph nodes metastases on long-term oncological outcomes; 3] to evaluate the current watch-and-wait (W&W) follow-up schedule and to propose improvements to make the follow-up schedule more efficient; 4] to evaluate the oncological and functional outcomes of a W&W approach in older patients; 5] to evaluate if distant metastasis occur later in W&W patients than in patients treated with CRT and total mesorectal excision (TME) by comparing metastasis and detection; and 6] to give an overview of current and new imaging technologies for prediction and assessment of response. The introduction *(chapter 1)* summarizes the currently available evidence and status with regard to these aims.

Part I: Patient selection and follow-up

Chapter 2 evaluates whether radiologists with variable levels of expertise are able to correctly identify those patients with substantial residual disease who require immediate surgery using a simplified three-categorized MRI response evaluation system. This may facilitate more selective use of endoscopy which is particularly interesting for low volume centres with limited access to endoscopy. Although there was not a perfect agreement between readers, radiologists correctly identified the 20% of poor responders who have substantial residual disease at histopathology and who require surgery without the need for endoscopy to evaluate response. Also, almost all complete responders were appointed to the good or intermediate response group. In total, by this approach 80% of the patients are referred for additional endoscopy to evaluate luminal response and decide on their eligibility for organ preserving therapy.

Chapter 3 describes a large pooled analysis of individual patient data from historical cohorts that evaluated the prevalence of lymph node metastases according to ypT-stage in patients with locally advanced rectal cancer treated with CRT and TME. Similar as in the primary setting, the risk of lymph node metastases increases with increasing T-stage: 7% for ypT0, 12% for ypT1, 17% for ypT2, 40% for ypT3 and 46% for ypT4. In addition, the presence of malignant lymph nodes was a strong predictor for poor long-term disease-free survival and overall survival in patients with ypT0-2 disease (HRs of 2.05-2.45) and in the total cohort (HRs of 2.08-2.26). These outcomes can be used in clinical practice to discuss the risk of lymph node metastases after local excision or when considering organ preservation according to the depth of residual tumour in the rectal wall.

Chapter 4 proposes a more efficient follow-up schedule by constructing a theoretical comparison of the occurrence and detection of local regrowth in the current follow-up schedule with four other hypothetical follow-up schedules. The new proposed follow-

up schedule reduces the number of examinations from 24 to 20. Endoscopy should be performed more frequently in the first two years because most regrowths occur within two years (98%), are located in the lumen (94%) and are visible on endoscopy (88%). After two years the schedule can be deintensified to yearly follow-up with endoscopy and MRI. It is likely that the proposed follow-up schedule will result in better efficiency and lower burden for patients.

Part II: Oncological outcomes

Chapter 5 evaluates the outcome of a W&W policy in older patients. W&W appears to be a safe alternative in older patients with a non-regrowth disease-free survival of 91%, overall survival of 97% and overall pelvic control of 98%. In addition, most patients avoided surgery and a permanent stoma with a colostomy-free rate of 93% and had reasonable anorectal and urinary function with good continence, no or minor Low Anterior Resection Syndrome Score and moderate urinary problems.

Chapter 6 compares the time pattern of distant metastasis from two large datasets of individual patient data: one dataset with 1642 W&W patients from the International Watch-and-Wait Database and the other dataset containing 2401 patients treated with CRT and surgery. The risk to develop metastases in W&W patients is low (3-year distant metastasis-free rate 92%) compared to patients treated with CRT and surgery (3-year distant metastasis-free rate 82%). Even though during the entire follow-up period W&W patients have a lower risk for distant metastases compared to operated patients, a significant interaction effect was found for follow-up time and distant metastasis: the risk for distant metastases was lower during early follow-up (HR of 0.27, 95%CI 0.12-0.60) than during later follow-up (HR 0.66, 95%CI 0.53-0.83) in W&W patients. So, although the risk to develop distant metastasis in W&W patients is low, they develop metastases later and the hypothesis is that there is a small risk that metastases originate from local regrowths.

Part III: Recent advanced imaging technologies

Chapter 7 gives an overview of current MR imaging techniques for response prediction and assessment of response and gives insight in which advanced imaging techniques will probably be valuable in future research and clinical practice. New functional imaging biomarkers (i.e. derived from diffusion weighted imaging or dynamic contrast enhanced MRI) capture changes in tumour perfusion and microstructure before they appear on morphological imaging. Many studies show that the combination of functional and morphological data has the highest accuracy in prediction and assessment of response, with other functional imaging biomarkers on the horizon. A recent promising advance is radiomics, which uses computer algorithms to find associations between quantitative features (i.e. pre-defined general purpose features such as intensity distribution, tumour shape and heterogeneity) and clinical outcome. Several studies have evaluated radiomics based on MRI of rectal tumours, with promising results in response prediction and post-neoadjuvant therapy assessment of response. However, due to various methodological

problems, the generalizability is hampered and, more studies are needed before it can be adopted into clinical practice.

Chapter 8 evaluates the accuracy of deep learning models based on endoscopic images acquired during response evaluation to identify complete responders. The diagnostic performance of different open access deep learning networks was evaluated. In addition, the diagnostic performance of clinical features were combined with endoscopic image features. The outcomes of the different models were compared with the reference standard: complete response (sustained complete response ≥2 years or ypT0 at histopathology) or residual disease (local regrowth) during follow-up or ypT1-4 at histopathology). Findings show that deep learning has a modest accuracy to detect complete responders and the combined models achieved the highest performance (AUC 0.76-0.83) compared to models that were built on clinical features (AUC 0.71-0.74) or endoscopic images only (AUC 0.71-0.79). Overall, EfficientNet-B2 achieved the highest performance amongst the different combined models with an AUC of 0.83 , an accuracy of 0.75, a sensitivity of 0.77 and specificity of 0.75.

APPENDICES

Nederlandse samenvatting

De doelen van dit proefschrift zijn: 1] evalueren of magnetic resonance imaging (MRI) als enige beeldvorming nauwkeurig patiënten kan identificeren die na chemoradiatie een substantiële tumorrest hebben waarvoor de standaard invasieve behandeling (TME) noodzakelijk is, en of deze bevindingen reproduceerbaar zijn onder radiologen met verschillende expertiseniveaus; 2] om de gepoolde prevalentie van persisterende maligne lymfeklieren na chemoradiatie te evalueren op basis van resterende tumordoorgroei in de rectale wand (ypT) en om te beoordelen wat de impact is van persisterende lymfekliermetastasen na chemoradiatie op de oncologische uitkomsten op de lange termijn; 3] het huidige watch-and-wait (W&W) follow-up schema evalueren en verbeteringen voorstellen om het follow-up schema efficiënter te maken om eerder lokale teruggroei van tumor (regrowth) te detecteren maar de belasting voor de patient zo laag mogelijk te houden; 4] om de oncologische en functionele resultaten van een W&W behandeling bij oudere patiënten te evalueren; 5] om te evalueren of afstandsmetastasen later optreden bij W&W patiënten dan bij patiënten die worden behandeld met chemoradiatie en TME; en 6] om een overzicht te geven van de huidige en nieuwe beeldvormingstechnieken ten aanzien van het voorspellen en beoordelen van de tumorrespons en handreikingen voor de toekomst te geven. In de inleiding (*hoofdstuk 1*) worden de momenteel beschikbare resultaten met betrekking tot deze doelstellingen samengevat.

Deel I: Patiëntselectie en follow-up

In *hoofdstuk 2* wordt geëvalueerd of radiologen met verschillende expertiseniveaus met behulp van een vereenvoudigd MRI-responsevaluatiesysteem (goede, redelijke of slechte respons) adequaat patiënten kunnen identificeren die na chemoradiatie een substantiële tumorrest hebben en gebaat zijn bij een TME. Een dergelijke benadering is vooral interessant voor laag volume centra met beperkte endoscopie mogelijkheden, waardoor er selectiever gebruik gemaakt kan worden van endoscopie ter beoordeling van de respons. Ondanks dat er geen perfecte overeenstemming was tussen de verschillende radiologen, werden de 20% patiënten door alle radiologen in de studie met een histopathologisch bewezen resttumor adequaat geïdentificeerd als slechte respons. In die gevallen is een operatieve ingreep geïndiceerd en zou er dus terecht worden afgezien van aanvullende endoscopie. Ook werden bijna alle complete responders adequaat toegewezen aan de groep met een goede of redelijke respons. Middels deze methode zou 80% van de patiënten worden doorverwezen voor aanvullende endoscopie ter beoordeling van de luminale respons en om te beslissen of ze in aanmerking zouden komen voor orgaansparende therapie.

Hoofdstuk 3 beschrijft een grote gepoolde analyse van individuele patiëntgegevens uit historische cohorten die de prevalentie van lymfekliermetastasen evalueerden op basis van het ypT-stadium bij patiënten met lokaal gevorderde rectumcarcinoom die waren behandeld met chemoradiatie en TME. Net als in de primaire setting neemt het risico op lymfekliermetastasen toe bij hogere tumorstadia: 7% voor ypT0, 12% voor ypT1, 17% voor ypT2, 40% voor ypT3 en 46% voor ypT4. Bovendien was de aanwezigheid van maligne lymfeklieren een sterke voorspeller voor een slechte ziektevrije overleving en slechte totale

overleving op de lange termijn bij patiënten met een ypT0-2 tumor (HRs van 2.05-2.45) en in het totale cohort (ypT0-4, HRs van 2.08-2.26). Deze uitkomsten kunnen worden gebruikt in de klinische praktijk voor het bespreken van de risico's op lymfekliermetastasen op basis van de resterende tumordoorgroei in de rectale wand na lokale excisie of bij watchful waiting na chemoradiatie.

In *hoofdstuk 4* wordt een voorstel gedaan om het huidige W&W follow-up schema efficiënter te maken. Hiervoor wordt een theoretische vergelijking uitgevoerd waarin het optreden en detecteren van regrowth in het huidige follow-up schema vergeleken wordt met vier andere hypothetische follow-up schema's. Het nieuwe voorgestelde follow-up schema vermindert het totaal aantal aanvullende onderzoeken in 5 jaar van 24 naar 20. Hierbij vinden gedurende de eerste twee jaar frequenter endoscopieën plaats, omdat de meeste regrowths binnen twee jaar optreden (98%), zich met name in het lumen (94%) bevinden en vrijwel altijd zichtbaar zijn bij endoscopie (88%). Na twee jaar kan het schema worden afgeschaald tot jaarlijkse follow-up met endoscopie en MRI. Het is waarschijnlijk dat het voorgestelde follow-up schema efficiënter en minder belastend zal zijn voor patiënten.

Deel II: Oncologische uitkomsten

Hoofdstuk 5 evalueert de uitkomsten van het W&W beleid bij oudere patiënten (≥75 jaar). W&W lijkt in deze groep een veilig alternatief te zijn met een totale overleving van 97%, een niet-regrowth ziekte vrije overleving van 91% (het uitblijven van afstandsmetastasen of overlijden) en een pelvic control van 98% (het uitblijven van een locoregionaal recidief in het bekken). Bovendien kon bij de meeste patiënten chirurgie en een permanent stoma vermeden worden, met een stomavrij percentage van 93%, en hadden patiënten een redelijke anorectale functie en weinig mictieproblemen.

In *hoofdstuk 6* wordt het tijdspatroon van het ontwikkelen van afstandsmetastasen vergeleken in twee grote datasets van individuele patiëntgegevens: een dataset met 1642 W&W patiënten uit de International Watch-and-Wait Database en de andere dataset met 2401 patiënten behandeld met chemoradiatie en TME. Het risico op het ontwikkelen van metastasen bij W&W patiënten is laag (3-jaar metastasevrij percentage van 92%) vergeleken met patiënten behandeld met chemoradiatie en TME (3-jaar metastasevrij percentage van 82%). Hoewel W&W patiënten gedurende de gehele follow-up periode een lager risico hebben op het ontwikkelen van afstandsmetastasen in vergelijking met geopereerde patiënten, werd een significant interactie-effect gevonden voor de duur van follow-up en metastasen op afstand: het risico op het ontwikkelen van afstandsmetastasen was lager in de vroege follow-up periode (HR 0.27, 95%CI 0.12-0.60) vergeleken met de latere follow-up periode (HR 0.66, 95%CI 0.53-0.83) in W&W patiënten. Dus, hoewel het risico op het ontwikkelen van afstandsmetastasen bij W&W patiënten laag is, ontwikkelen metastasen zich mogelijk later met de hypothese dat er een klein risico is dat de metastasen zich ontwikkelen door het optreden van regrowths.

Deel III: Recente geavanceerde beeldvormingstechnieken

Hoofdstuk 7 geeft een overzicht van het voorspellen en beoordelen van de respons middels huidige MR-beeldvormingstechnieken en geeft inzicht in welke geavanceerde beeldvormingstechnieken potentieel waardevol zullen zijn voor toekomstig onderzoek en de klinische praktijk. Nieuwe functionele beeldvorming die gebruik maakt van biomarkers (d.w.z. afgeleid van diffusion-weighted imaging of dynamic contrast enhanced MRI) detecteren veranderingen in tumorperfusie en microstructuur voordat dit zichtbaar is op morfologische beeldvorming (T2- en T1-gewogen MRI of CT). Onderzoek heeft aangetoond dat de combinatie van functionele en morfologische gegevens de hoogste nauwkeurigheid heeft ten aanzien van de voorspelling en beoordeling van de respons. Daarbij zijn aanvullende functionele beeldvormende biomarkers in opmars. Een voorbeeld hiervan is radiomics, dat door middel van computeralgoritmen associaties vindt tussen kwantitatieve kenmerken (d.w.z. vooraf gedefinieerde tumorkarakteristieken op beeldvorming afgeleid van vorm, relatie met omgeving, interne structuur, intensiteit en densiteit etc.) en de klinische uitkomstmaten. Verschillende studies die zijn gebaseerd op het gebruik van MRI bij patiënten met rectumcarcinoom laten positieve resultaten zien over het gebruik van radiomics voor responspredictie en beoordeling van de respons na neoadjuvante chemoradiatie. Een ander voorbeeld van kunstmatige intelligentie is deep learning waarbij de computer automatisch grote hoeveelheden data kan analyseren en hierbij sneller is dan de mens. Echter, door verschillende methodologische problemen zijn de resultaten nog niet te generaliseren naar de praktijk en zijn er meer studies nodig voordat het in de klinische praktijk kan worden toegepast.

In *hoofdstuk 8* wordt de diagnostische accuratesse van het gebruik van deep learning modellen op endoscopiebeelden die zijn verkregen tijdens responsevaluatie geëvalueerd om patiënten met een complete respons te identificeren. De accuratesse van deep learning werd geëvalueerd door middel van het gebruik van open access deep learning modellen waarbij de uitkomsten werden gecombineerd met klinische kenmerken en endoscopische beeldkenmerken. De uitkomsten van de verschillende modellen werden vervolgens vergeleken met de gouden standaard: complete respons (aanhoudende complete respons ≥2 jaar tijdens watchful waiting of ypT0 bij histopathologie) of resttumor (regrowth) tijdens follow-up of ypT1-4 bij histopathologie). De uitkomsten laten zien dat deep learning redelijk nauwkeurig complete responders kan detecteren, waarbij de gecombineerde modellen (klinische factoren en beeldkarakteristieken) enigszins nauwkeuriger waren (AUC 0.76-0.83) in vergelijking met modellen die waren gebaseerd op alleen klinische kenmerken (AUC 0.71-0.74) of alleen endoscopische beeldkenmerken (AUC 0.71-0.79). Van de verschillende gecombineerde modellen had EfficientNet-B2 de beste uitkomsten met een AUC van 0.83, een accuratesse van 0.75, een sensitiviteit van 0.77 en een specificiteit van 0.75. Deze uitkomsten laten zien dat deep learning mogelijk een rol kan spelen bij het beoordelen van de tumor response en dat meer onderzoek nodig is om dit verder te evalueren.

APPENDICES

Impact

In this chapter, the scientific impact of the results described in this thesis will be outlined, as well as the social impact anticipated or already achieved by discussing the following four aspects:
1. Research
2. Relevance
3. Target group
4. Activity

Research

The main objective of the research described in this thesis was to provide an overview of the challenges in organ preservation in patients with rectal cancer and provide new data regarding patient selection, follow-up and outcomes. The most important results and conclusions of this thesis are: [1] a simplified three-categorized MRI response evaluation system aids radiologists with variable levels of expertise to identify patients who have substantial residual disease who require immediate surgery rather than further response assessment; [2] deep learning based on endoscopy images after chemoradiation (CRT) has a modest accuracy to detect complete responders; [3] the risk of lymph node metastasis increases with increasing depth of residual disease in the tumour wall after CRT, similar to the setting of total mesorectal excision (TME) without neoadjuvant therapy, and the presence of remaining malignant lymph node metastases is a strong predictor for poor outcome, independent from the ypT-stage; [4] the efficiency of the watch-and-wait (W&W) follow-up schedule increases when follow-up is intensified in the first two years and deintensified after two years; [5] the oncological and functional outcomes of older patients who follow a W&W approach are very good; [6] the risk of metastases in W&W patients is low, but there may be a small risk that some metastases originate from local regrowths; and [7] multiple MR imaging techniques are valuable for response prediction and response assessment and more techniques are on the horizon such as AI modelling.

Relevance

This thesis is relevant to clinical practice as it gives tools to finetune the selection and follow-up for W&W in complete responders after CRT. Furthermore, information is provided for counseling of the patient for a W&W policy in the outpatient clinic. Based on the data from this thesis physicians and patients are better informed on the potential risks and benefits of W&W in complete responders (including the elderly). This will facilitate shared decision-making. Last, this thesis guides future research with an emphasis on exploration of new techniques to enable more accurate response prediction and assessment in rectal cancer, with the ultimate goal to increase organ preservation rates without compromising oncological outcomes.

Target group

There are several people who could benefit from the results presented in this thesis. Firstly, in about 20% of locally advanced rectal cancer patients who are treated with neoadjuvant

CRT a clinical complete response is found and in these patients organ preserving treatment could be considered. In addition, patients with a good but not complete response (near complete response) after neoadjuvant treatment or those who are not suitable for surgery due to a high risk of morbidity and mortality might be candidates for organ preservation. Besides, there is a rising interest to aim for organ preservation by neoadjuvant treatment in patients who have a small rectal tumour instead of upfront surgery. These patients all benefit from the results of this thesis. Second, the results in this thesis are also interesting for the multidisciplinary team that deals with rectal cancer, for example physician assistants and clinicians such as surgeons, oncologists, gastroenterologists, radiologists and pathologists, in The Netherlands, but also abroad.

Activity

The website of the Netherlands Cancer Institute provides some more background information about W&W patients that are interested in W&W but who are not familiar with the approach.[1] Also, a former W&W patient created a website specifically about the W&W approach[2] and additionally wrote two books regarding his experience and those of others with W&W.[3, 4] This information is specifically interesting for patients with rectal cancer who are potentially eligible or considering to follow a W&W program. Through these channels patients can be informed about the results of this thesis and other results achieved by the W&W group.

The website of the Netherlands Cancer Institute might also be interesting for clinicians who are unfamiliar with W&W. In addition, every two or three years the Netherlands Cancer Institute organizes a national W&W symposium for expert centres and dedicated clinicians from the Netherlands. Every five years an international W&W symposium is being held in Lisbon organized by the International Watch-and-Wait Database (IWWD) in order to discuss the most up-to-date literature and future perspectives regarding organ preserving therapy. The knowledge gained during these meetings should be passed on to the different multidisciplinary teams of W&W expert centres and/or dedicated clinicians.

For a selected group of patients organ preservation has proven to be feasible and oncologically safe, and currently there is a focus on making organ preservation as a treatment option available to an increasing number of patients. At the same time there is a need for more data on functional outcome and quality of life. The current Dutch network originated in Maastricht University and has now been expanded and coordinated by The Netherlands Cancer Institute, with a connection to the IWWD. Maastricht University and The Netherlands Cancer Institute remain close cooperation partners in this network that provides a unique opportunity to set up new studies and provide more data. The Dutch Cancer Society has previously awarded grants to facilitate research in the field of organ preservation in rectal cancer, including the studies in the present thesis. As a result of all these efforts, The Netherlands has at present a leading position in this field, and further funding is required to retain this position.

REFERENCES

1. Wait-and-See bij rectumkanker. Antoni van Leeuwenhoek. Accessed 25-08-2021, https://www.avl.nl/alles-over-kanker/overzicht-van-alle-behandelingen/wait-and-see-bij-rectumkanker/
2. Keusters H. Accessed 25-08-2021, https://harriekeusters.nl/
3. Keusters H. *Nooduitgang: een bijzonder verhaal over endeldarmkanker*. 2015.
4. Keusters H. *Dom geluk, vette pech*. Scriptum Books; 2019.

APPENDICES

List of publications

Custers PA, Hupkens BJP, Grotenhuis BA, Kuhlmann KFD, Breukink SO, Beets GL, Melenhorst J; **Dutch Watch-and-Wait Consortium**. Selected stage IV rectal cancer patients managed by the watch-and-wait approach after pelvic radiotherapy: a good alternative to total mesorectal excision surgery? Colorectal Dis. 2022 Jan 21. doi: 10.1111/codi.16034. Online ahead of print.

Bahadoer RR, Peeters KCMJ, Beets GL, Figueiredo NL, Bastiaannet E, Vahrmeijer A, Temmink SJD, Meershoek-Klein Kranenbarg WME, Roodvoets AGH, Habr-Gama A, Perez RO, van de Velde CJH, Hilling DE; **International Watch & Wait Database Consortium (IWWD).** Watch and wait after a clinical complete response in rectal cancer patients younger than 50 years. Br J Surg. 2021 Dec 17;109(1):114-120.

Haak HE, Gao X, Maas M, Waktola S, Benson S, Beets-Tan RGH, Beets GL, van Leerdam M, Melenhorst J. The use of deep learning on endoscopic images to assess the response of rectal cancer after chemoradiation. Surg Endosc. 2021 Oct 12. doi: 10.1007/s00464-021-08685-7. Online ahead of print.

Haak HE, Beets GL, Peeters K, Nelemans PJ, Valentini V, Rödel C, Kuo L, Calvo FA, Garcia-Aguilar J, Glynne-Jones R, Pucciarelli S, Suarez J, Theodoropoulos G, Biondo S, Lambregts DMJ, Beets-Tan RGH, Maas M. Prevalence of nodal involvement in rectal cancer after chemoradiotherapy. Br J Surg. 2021 Oct 23;108(10):1251-1258.

Marres CCM, Engelmann EWM, Buskens CJ, **Haak HE**, Bemelman WA, van de Ven AWH. The importance of rectal contrast in CT assessment to detect anastomotic leakage after colorectal surgery. Colorectal Dis. 2021 Sep;23(9):2466-2471.

Haak HE, Žmuc J, Lambregts DMJ, Beets-Tan RGH, Melenhorst J, Beets GL, Maas M; Dutch Watch-and-Wait Consortium. The evaluation of follow-up strategies of watch-and-wait patients with a complete response after neoadjuvant therapy in rectal cancer. Colorectal Dis. 2021 Jul;23(7):1785-1792.

Haak HE, Maas M, Trebeschi S, Beets-Tan RGH. Modern MR imaging Technology in Rectal Cancer; There Is More than Meets the Eye. Front Oncol. 2020 Oct 8;10:537532.

van Ommeren-Olijve SJ, Burbach JPM, Furnée EJB; **Dutch Snapshot Research Group.** Risk factors for non-closure of an intended temporary defunctioning stoma after emergency resection of left-sided obstructive colon cancer. Int J Colorectal Dis. 2020 Jun;35(6):1087-1093. doi: 10.1007/s00384-020-03559-1.

Haak HE, Beets GL. Evaluation of Tumour Response After Radiotherapy in Rectal Cancer. Multidisciplinary Treatment of Colorectal Cancer. Springer, Cham. Chapter 26, 2020. doi:10.1007/978-3-030-58846-5_26

Haak HE, Maas M, Lahaye MJ, Boellaard TN, Delli Pizzi A, Mihl C, van der Zee D, Fabris C, van der Sande ME, Melenhorst J, Beets-Tan RGH, Beets GL, Lambregts DMJ. Selection of Patients for Organ Preservation After Chemoradiotherapy: MRI Identifies Poor Responders Who Can Go Straight to Surgery. Ann Surg Oncol. 2020 Mar 14. doi:10.1245/s10434-020-08334-8.

Veld JV, Wisselink DD, Amelung FJ, Consten ECJ, de Wilt JHW, de Hingh I, Bemelman WA, van Hooft JE, Tanis PJ; **Dutch Snapshot Research Group.** Synchronous and Metachronous Peritoneal Metastases in Patients with Left-Sided Obstructive Colon Cancer. Ann Surg Oncol. 2020 Mar 13. doi:10.1245/s10434-020-08327-7.

Haak HE, Maas M, Lambregts DMJ, Beets-Tan RGH, Beets GL; Dutch Watch-and-Wait Consortium. Is watch and wait a safe and effective way to treat rectal cancer in older patients? Eur J Surg Oncol. 2020 Mar;46(3):358-362. doi: 10.1016/j.ejso.2020.01.005.

Amelung FJ, Borstlap WAA, Consten ECJ, Veld JV, van Halsema EE, Bemelman WA, Siersema PD, Ter Borg F, van Hooft JE, Tanis PJ; **Dutch Snapshot Research Group**. Propensity score-matched analysis of oncological outcome between stent as bridge to surgery and emergency resection in patients with malignant left-sided colonic obstruction. Br J Surg. 2019 May 10. doi: 10.1002/bjs.11172.

APPENDICES

Dankwoord

Met veel plezier heb ik 3 jaar mogen werken in het Antoni van Leeuwenhoek ziekenhuis. Dit proefschrift had er niet mogen zijn zonder hulp en steun van velen.

Mijn **promotoren**, allereerst **prof. Beets**, beste **Geerard**. Ondanks je overvolle agenda had je gedurende mijn promotietraject altijd tijd voor overleg. Daarnaast heb ik je nooit kunnen betrappen op enige vorm van stress of een slecht humeur. Ik bewonder je scherpe feedback op manuscripten en je kritische vragen tijdens congressen of bijeenkomsten van uiteenlopende onderwerpen. Ik heb je begeleiding als erg prettig ervaren en ik wil je hier graag voor bedanken.

Prof. Beets-Tan, beste **Regina**, je ambities en gedrevenheid voor de radiologie zijn bewonderenswaardig. Ondanks je nog drukkere agenda dan Geerard, maakte je altijd tijd voor mij vrij. Ook wil ik je graag bedanken voor het vertrouwen en je hulp voor de opleiding Radiologie.

Mijn **copromotor Dr. Maas**, beste **Monique**, ondanks dat ik gedurende de eerste periode van mijn onderzoek zonder jou moest beginnen, hebben wij een fijne band ontwikkeld. Op onderzoeksgebied benijd ik de manier hoe snel jij kan schakelen en direct de vinger op de zere plek kan leggen. Jouw schrijfmotto is: kort, korter, kortst, met als resultaat heldere en bondige manuscripten. Daarbij hield jij van strakke deadlines met als resultaat dat mijn promotie een flinke vaart had. Ook ben je ontzettend goed in statistiek, en was de hulp van een statisticus meermalen overbodig. Naast het onderzoek, werd er ook veel gepraat over andere dingen en heb ik met jou onder andere kunnen hebben over mijn toekomst bij de Radiologie. Ik wil je graag bedanken voor deze mooie periode en je goede adviezen.

Geachte leden van de promotiecommissie, **prof. Rutten**, **prof. de Bruïne**, **prof. Marijnen** en **dr. Horsthuis**, hartelijk dank voor het beoordelen van mijn proefschrift.

Beste leden van het W&W consortium, hartelijk dank voor jullie bijdrage en begeleiding van de wait and see studie. Beste **Dr. Melenhorst**, beste **Jarno**, veel dank voor je feedback en begeleiding van meerdere manuscripten.

Leden OOA-commissie: Johanna, **Marjanka** en **Max**, hartelijk dank voor jullie betrokkenheid en goede adviezen gedurende mijn promotietraject.

Lieve **paranimfen.** Lieve **Mathilde**, het was een welkome verrassing om een bekend gezicht uit Eindhoven te zien in het O-gebouw. Niet alleen op de kamer luisterend naar de non-stop (niet altijd even favoriete) “carwash” muzieklijst van Willem. Maar ook tijdens de vrijdagmiddag borrels die soms al op maandag middag begonnen, après-ski feesten in meerdere skioorden en de vele etentjes/koffietjes als afleiding van het phd-leven. Ik heb genoten van je verschillende Brabantse imitaties van onder andere Theo Maassen, chicken wings en internetgekkies. Bovenal ben ik heel blij dat je mijn paranimf wilt zijn. Lieve

Rebecca, het is altijd een verademing om met jou, naast vele andere dingen, ook te kunnen sparren over de medische wereld. Je staat altijd voor mij klaar en begrijpt de perikelen van arts zijn als geen ander. Ik ben ontzettend trots op jou hoe je alles voor elkaar hebt gebokst en dat je nu helemaal je plek hebt gevonden. Daarnaast ben ik ook ontzettend blij met jou als paranimf aan mijn zijde.

Alle **collega-onderzoekers van het Tuinhuis**, veel dank voor de fijne samenwerking. Beste **Doenja**, hartelijk dank voor je benaderbaarheid met het meermaals beoordelen van de MRI's van de W&W-studie, maar ook je visie en inbreng als medeauteur. Daarnaast wil ik jou ook graag bedanken voor je vertrouwen en hulp met de opleiding Radiologie.

Lieve (blonde) **W&S-onderzoekers**. Lieve **Marit**, dank voor je hulp tijdens de beginperiode van mijn onderzoek, we hebben mooie tijden gehad in Nice (samen met Britt) en Lissabon (met Petra)! Lieve **Petra**, veel dank voor de prettige samenwerking, eerst als begeleider bij je master scriptie en nadien als collega-onderzoeker en kamergenoot. Jouw nuchterheid en positiviteit zijn aanstekelijk. Lieve **Barbara**, heel veel succes met het vervolg van het W&S-onderzoek!

Lieve **Kete**, naast oud-huisgenoten werden wij ook nog eens collega's. Wat een feest! De vele koffietjes, wandelingen en lunchafspraken waren altijd een meer dan welkome afleiding. Niet alleen om te praten over het onderzoek maar juist ook om dat even lekker helemaal niet te doen. Het AvL heeft ons nader naar elkaar toe gebracht en daar ben ik ontzettend blij mee. Lieve **Emma**. Dank voor je luisterende oor, eerlijkheid, (kleding)adviezen en de gezellige etentjes met Mat. Kamer 13 was een hele fijne uitvalsbasis. Lieve **Maart** en **Wim**, aka SKICIE JODELO. Samen hebben wij de legendarische ski reis in het prachtige Schladming mogen organiseren. Ondanks dit nogal onbekende skioord was het met behulp van on-point Excel sheets met tijdsplanning tot op het kwartier, originele activiteiten (rodelend de rode piste af) en af en toe wat ouderlijk advies een zeer geslaagde week en vlogen de complimenten ons om de oren. Ik heb ontzettend van jullie genoten, en een ding is zeker: zonder ons, natte sokken! **Overige collega's van het O**, veel dank voor de mooie tijden tijdens de vrijdagmiddagborrels, wintersporten, weekendjes weg en OOA-retreat. Naast dat we natuurlijk ook hard aan het werk waren, was het ook zeker een tweede studententijd!

Al mijn **collega's binnen de afdeling Radiologie van het Amsterdam UMC**, hartelijk dank voor de ondersteuning in de laatste fase van de afronding van dit proefschrift, en ik kijk uit naar de komende jaren van de opleiding!

Lieve vrienden die een welkome afleiding waren en (onbewust) veel hebben bijgedragen aan dit boekje. Lieve **Eli**, door jou werd het zaadje gepland om te gaan promoveren, dus zonder jou was dit boekje er zeker niet gekomen! Lieve **Bar**, veel dank voor het immense scherm dat wij van jullie mochten overnemen, hierdoor werd thuiswerken veel aangenamer! Lieve **overige ex-roomies van het Lllorienhuis,** met name **Jive, Wokkel en Noor**. Ondanks

dat wij niet meer samenwonen zien wij elkaar gelukkig nog veelvuldig, en dat voelt altijd net als thuiskomen. Lieve **Olijf en Jet**. Dank voor alle fijne middagen en avonden. Lief **Eindje,** lieve **Tineke, Lieke, Katrien en Pascal.** Het blijft altijd een feest als we bij elkaar zijn en ik kijk ontzettend uit naar onze volgende mijlpalen. Lieve **Ghis**, zullen wij afspreken dat ik mij beter zal voelen op de dag van mijn verdediging dan tijdens mijn buluitreiking? Lieve **Olga**. Ik ben blij dat **Emilie** heeft bedacht dat wij vrienden moesten worden! Na eerst een aantal jaren samen coschappen te hebben gelopen, hebben we nu beiden onze eigen weg gevonden waarbij jij op dit moment aan het shinen bent in het buitenland als tropenarts, ontzettend knap!

Lieve **(schoon)familie**, lieve **Hans**, **Ellen**, **Laura**, **Jesper**, **Sanne**, **Mark**, **Diede** en **Teun**, dank voor jullie steun en betrokkenheid. Da ge bedankt zijt da witte he kul!

Lieve **oma Luc,** ondanks dat je in een andere tijd geboren bent, ben jij voor mij een volwaardig journalist. Naast dat je nog colleges volgt en elke krant van kaft tot kaft leest, ben je ook nog eens op de hoogte van minder belangrijke maar zeker niet minder interessante dingen. Ik geniet nog altijd van je kritische vragen of krantenknipsels en boeken die ik onder mijn neus geschoven krijg als ik bij je op bezoek kom. Ik ben ontzettend blij om jouw kleindochter te mogen zijn, en ik hoop (en weet zeker) op nog vele jaren samen!

Lieve **Bas**, je bent een broer om tegen op te kijken, niet alleen als onderzoeker of arts, maar ook als allesweter van elke (on)belangrijk feitje. Ik vind het ontzettend knap wat jij allemaal uit je dag kan halen, en elke keer weer het maximale voor elkaar krijgt. Ik geniet van onze wandelingen samen gepaard met koffie en je luisterende oor. Ook ben ik heel blij dat jij en Megan voor nu nog even in Nederland blijven, zodat we elkaar nog veelvuldig kunnen zien. Ik heb veel zin in aankomend jaar om alle mooie dingen die op de planning staan samen te vieren.

Lieve **mama** en **papa**, jullie zijn beide harde werkers met veel ambitie en een onuitputtelijk doorzettingsvermogen. Jullie onvoorwaardelijke steun en vertrouwen naar elkaar vind ik ontzettend dierbaar. Maar bovenal zijn jullie de allerfijnste en warmste ouders die ik mijzelf kan wensen. Wat er ook gebeurd, en waar ik ook ben, jullie staan altijd voor mij klaar. Ik ben heel blij met jullie en ik hou ontzettend veel van jullie.

Lieve **Freek,** je enthousiasme, positiviteit en humor helpen mij de dagen door. Je staat met beide benen op de grond en bent een koning in relativeren. De tijd die wij samen als collega-huisgenoten hebben doorgebracht vanwege COVID was een cadeautje. Niet alleen was het fijn om een extra hulplijn te hebben over onderzoekdingen, maar ook om de soms nogal saaie promotiedagen kleur te geven met wandelingen-naar-werk-koffie en de "get ready for the launch" deunen door de speakers tijdens lunchtijd. Ik ben ontzettend blij met jou en ik heb heel veel zin in de mooie tijd die we samen gaan beleven. Ik hou van jou.

APPENDICES

Curriculum vitae

Hester Haak was born on February 22nd, 1991, in 's Gravenhage, the Netherlands. She graduated from secondary school at the Lorentz-Casimir Lyceum in Eindhoven, the Netherlands in 2009. After graduation, she started her medical training at the University of Amsterdam in 2009, with additional research internships at the Department of Surgery at the Royal London Hospital and the Blizard Institute in London, the United Kingdom. Afterwards, she received her medical degree at the University of Amsterdam in 2016. Following this year, she worked as a surgical resident not in training at the Department of Surgery at Flevoziekenhuis, Almere, the Netherlands. In 2018, Hester started as a PhD candidate at Maastricht University (GROW) and the Netherlands Cancer Institute at the Department of Surgery and Radiology under the supervision of prof. dr. Beets, prof. dr. Beets-Tan and dr. Maas. She focused on the watch-and-wait policy for complete responders after chemoradiotherapy for rectal cancer. During her PhD, Hester coordinated the national multicentre watch-and-wait implementation study in 15 hospitals. In addition, she presented at many national and international conferences and got awarded for "top 5 best presentations" at the European Society of Surgical Oncology conference (ESSO) in 2019. She was invited to give a presentation in Boston, Massachusetts, the USA for which she received a travel grant. In April 2021 she started her residency program Radiology and Nuclear Medicine at the Amsterdam University Medical Centres, location VUmc.