This is a repository copy of Phase 2 Neoadjuvant Treatment Intensification Trials in Rectal Cancer: A Systematic Review.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/122035/

Version: Accepted Version

Article:
Teo, MTW, McParland, L, Appelt, AL orcid.org/0000-0003-2792-9218 et al. (1 more author) (2018) Phase 2 Neoadjuvant Treatment Intensification Trials in Rectal Cancer: A Systematic Review. International Journal of Radiation Oncology Biology Physics, 100 (1). pp. 146-158. ISSN 0360-3016

https://doi.org/10.1016/j.ijrobp.2017.09.042

© 2017, Elsevier Inc. This manuscript version is made available under the CC BY-NC-ND 4.0 license https://creativecommons.org/licenses/by-nc-nd/4.0/

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can’t change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Accepted Manuscript

Phase II neoadjuvant treatment intensification trials in rectal cancer: a systematic review

Mark T.W. Teo, FRCR, Lucy McParland, MSc, Ane L. Appelt, PhD, David Sebag-Montefiore, FRCR

PII: S0360-3016(17)33948-2
DOI: 10.1016/j.ijrobp.2017.09.042
Reference: ROB 24521

To appear in: International Journal of Radiation Oncology • Biology • Physics

Received Date: 11 April 2017
Revised Date: 23 August 2017
Accepted Date: 21 September 2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Title: Phase II neoadjuvant treatment intensification trials in rectal cancer: a systematic review.

Authors: Mark T.W. Teo FRCR\textsuperscript{1,2}, Lucy McParland MSc\textsuperscript{1,3}, Ane L. Appelt PhD\textsuperscript{1,2}, David Sebag-Montefiore FRCR\textsuperscript{1,2}

Affiliations:
\textsuperscript{1} Radiotherapy Research Group, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, United Kingdom
\textsuperscript{2} Leeds Cancer Centre, St James University Hospital, Leeds, United Kingdom
\textsuperscript{3} Clinical Trials Research Unit, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, United Kingdom

Shortened Title: Rectal cancer neoadjuvant phase II trials

Corresponding author: Prof D Sebag-Montefiore
Radiotherapy Research Group, Level 4 Bexley Wing, St James University Hospital, Beckett Street, Leeds LS9 7TF, United Kingdom
Tel No: +44 (0)113 206 8586
Fax No: +44 (0)113 206 8474
E-mail: d.sebag-montefiore@leeds.ac.uk

Authors responsible for statistical analyses:
Dr Mark TW Teo
Radiotherapy Research Group, Level 4 Bexley Wing, St James University Hospital, Beckett Street, Leeds LS9 7TF, United Kingdom
Tel No: +44 (0)113 206 7614
E-mail: m.t.w.teo@leeds.ac.uk
Ms Lucy McParland

Radiotherapy Research Group, Level 4 Bexley Wing, St James University Hospital, Beckett Street, Leeds LS9 7TF, United Kingdom

Tel No: +44 (0)113 343 9050
E-mail: L.McParland@leeds.ac.uk

Conflict of Interest Notification: None.

Acknowledgements:

AA is supported by Yorkshire Cancer Research Academic Fellowship funding (grant L389AA)
Abstract

Purpose

Multiple phase II trials of neoadjuvant treatment intensification in locally advanced rectal cancer have reported promising efficacy signals but these have not translated into improved cancer outcomes in phase III trials. Improvements in phase II trial design are needed to reduce these false positive signals. This systematic review evaluated the design of phase II trials of neoadjuvant long-course (chemo)radiotherapy treatment intensification in locally advanced rectal cancer.

Methods and Materials

PubMed, EMBASE, MEDLINE and Cochrane Library were searched for published phase II trials of neoadjuvant treatment intensification from 2004-2016. Trial clinical design and outcomes were assessed, with statistical design and compliance rated using a previously published system. Multivariable meta-regression analysis of pathological complete response (pCR) was conducted.

Results

Ninety-two eligible trials were identified. Patients with AJCC stage II and III equivalent disease were eligible in 87(94.6%) trials. Forty-three(46.7%) trials mandated MRI local staging. Only 12(13.0%) trials were randomised, with eight having a standard treatment control arm. Just 51(55.4%) trials described their statistical design with 21(22.8%) trials failing to report their sample size derivation. The majority of trials (n=84, 91.3%) defined a primary endpoint but 15 different primary endpoints were used. All trials reported pCR rates. Only 38(41.3%) trials adequately reported trial statistical design and compliance. Meta-analysis revealed a pooled pCR rate of 17.5%(95%CI:15.7%-19.4%) across treatment arms of neoadjuvant long-course (chemo)radiotherapy treatment intensification and
substantial heterogeneity amongst the reported effect sizes ($I^2=55.3\%$, $P<0.001$). Multivariable meta-regression analysis suggested increased pCR rates with higher radiotherapy doses (adjusted $p=0.025$).

**Conclusions**

Improvement in the design of future phase II rectal cancer trials is urgently required. A significant increase in randomised trials is essential to overcome selection bias and determine novel schedules suitable for phase III testing. This systematic review provides key recommendations to guide future treatment intensification trial design in rectal cancer.
Introduction

In rectal cancer, total mesorectal excision (TME) remains the primary curative treatment. In operable rectal cancer, phase III randomised trials have demonstrated improved local control with neoadjuvant concomitant fluoropyrimidine long-course chemoradiotherapy[1, 2] and short-course radiotherapy[3, 4]. Overall 5-year local recurrence and disease-free survival rates of 4.7-7.0% and 68.0-73.6% respectively are reported, with more recurrences in Stage III disease[2, 3, 5].

Phase II trials play a significant role in identifying experimental regimens showing early efficacy signals and acceptable toxicity warranting further testing in definitive phase III trials. In the past decade, many phase II trials have focused on neoadjuvant treatment intensification of long-course (chemo)radiotherapy in rectal cancer, by radiotherapy dose escalation, combination chemotherapy regimens, or the addition of biological agents or putative radiosensitisers[6]. So far, promising efficacy signals in these phase II trials have not translated into improved cancer outcome in phase III trials. Of six phase III trials[7-12] testing the addition of concurrent oxaliplatin to long-course (chemo)radiotherapy, only one has reported an improvement in disease-free survival in peer-reviewed publication. However, the trials reported a 3-21% absolute increase in grade 3/4 toxicity.

This systemic review aims to assess the clinical and statistical design of phase II rectal cancer trials of neoadjuvant long-course (chemo)radiotherapy treatment intensification to better understand where improvements can be made. A meta-regression analysis explored the effect of clinical design factors on the commonly used early efficacy endpoint of pathological complete response (pCR).
Methods and Materials

Search strategy

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) group guidelines[13].

Search terms included all synonyms for phase II trials, rectal cancer, neoadjuvant (chemo)radiotherapy and response outcomes. Further details are provided in the Appendix. The PubMed, EMBASE, MEDLINE and the Cochrane Library databases were searched. Retrieved articles reference lists were reviewed for additional articles.

Eligible trials were defined as peer-reviewed publications of phase II trials in rectal adenocarcinoma involving neoadjuvant long-course (chemo)radiotherapy treatment intensification (either radiotherapy dose, additional chemotherapy or biological agents, or radiosensitisers) with reported outcome measures (i.e. pathological response, survival or toxicity); written in the English language, between 1st January 2004 and 3rd August 2016 (the last search date). Updated data in later publications of the same cohort were included. Long-course radiotherapy was defined as radiotherapy treatment time of at least two weeks and with a radiotherapy equivalent dose in 2Gy fractions (EQD2) of 40Gy or greater. Exclusion criteria were conference abstracts; retrospective series, case reports or epidemiological trials; phase I or phase I/II trials; radiotherapy EQD2 less than 40Gy; and recurrent or metastatic disease. Trials with reduced radiotherapy doses, including the use of short-course radiotherapy schedules, were excluded to try to reduce trial treatment heterogeneity with a focus on the intensification of chemo-radiotherapy treatment.
Data extraction and full text analysis

Trial eligibility criteria, treatment protocols, statistical design, outcomes, and the authors’ overall conclusion of the trial results were recorded for each publication where available. The authors’ final conclusion was subjectively inferred to be negative by the Reviewers (MT and LM) if the authors concluded that the experimental regimen was detrimental or had no added benefit; neutral if no conclusion was stated; and positive if the agent was concluded to be efficacious with acceptable toxicity or worthy of further investigation. The Reviewer’s critical appraisal of the statistical design, protocol compliance and interpretation of results was recorded (as defined by Mariani & Marubini[14]). The first Reviewer evaluated all articles, with 20% of articles verified by the second Reviewer.

Statistical methods

Standard treatment dosing

As there remains no international consensus on standard chemoradiotherapy dosing, “standard” treatment dosing was based on control arm doses used in phase III trials opened prior to 2004[3, 5, 15] defined as a radiotherapy EQD2 dose of 44 to 54 Gy (α/β=11.1Gy[16]) with concurrent 5-fluorouracil (200-300mg/m\(^2\) daily, days of radiotherapy or weekly; or 350-500mg/m2 daily over 5 days, weeks 1 and 5 of radiotherapy), or capecitabine (≥1600mg/m\(^2\) taken daily or on days of radiotherapy) chemotherapy. EQD2 calculations included endorectal intracavitary brachytherapy boosts but excluded intraoperative radiotherapy boosts. For non-standard fractionation schedules, a tumour repopulation dose increment of 0.15Gy/day[16] was used relative to a conventional 25 fraction (5-week) course.
**Meta-analysis**

A meta-analysis was conducted to estimate the pooled pCR rate, across treatment arms of neoadjuvant long-course (chemo)radiotherapy treatment intensification. A meta-regression analysis was not performed for long-term outcomes due to variable reporting of these end-points. For this analysis, pCR was defined as ypT0ypN0 or if not stated, the trial-defined pCR. A random-effects model[17] was used to derive a pooled estimate of the pCR rate and 95% confidence intervals (CIs). Cochran’s Q-test and Higgin’s I² statistic were used to assess heterogeneity across reported effect sizes. A multivariable meta-regression analysis was used to explore the effect of pre-specified clinical design factors on the observed heterogeneity, and the pooled treatment effect estimate. The clinical design factors investigated included: single vs multi-centre studies, compromise of standard concurrent fluoropyrimidine chemotherapy dose (yes vs. no), recommended minimum interval from end of (chemo)radiotherapy to surgery (≥6 vs. <6 weeks), radiotherapy EQD2 dose, addition of a second concurrent chemotherapy agent (yes vs. no) and addition of chemotherapy before or after (chemo)radiotherapy (yes vs. no). Multiple testing was adjusted for using a permutation test[18]. All statistical analyses were conducted at the 5% significance level using Stata Release 13 (StataCorp LP, USA). For the primary analysis, multi-arm trials which were in the minority in this systematic review, were excluded to reduce the between trial heterogeneity. A sensitivity analysis was conducted to assess of the effect of including treatment intensification arms from multi-arm trials, including ‘study’ as a covariate to account for the correlation between treatment arms from the same study.
Results

General trial characteristics

Figure 1 summarises the systematic article search results and reasons for exclusion. A total of 92 phase II trials were included in this review[19-110] (Supplementary Table 1). Trials were predominantly single-centre studies (n=53, 57.6%) (Table 1). The median number of patients recruited per trial was 47 (range: 8-144 patients). The median recruitment period was 26.9 months (range: 4.0-104.0 months), excluding 12 trials not reporting their recruitment duration. Twenty-three (25%) trials reported registration on a clinical trials registry.

Eligibility criteria

Age inclusion criteria were reported in 74 (80.4%) of the 92 trials. A majority of trials accepted a WHO performance status (PS) of 0-2 (n=53, 57.6%).

The local disease staging inclusion criteria were equivalent to American Joint Committee on Cancer (AJCC) stage II/III disease for 87(94.6%) trials, with three trials limiting trial entry to AJCC stage III disease only (Table 1). One trial included patients with “non-metastatic disease at a locally advanced stage that made R0 resection and sphincter preservation uncertain”[61], and another trial enrolled patients with AJCC stage I-III disease[28].

Trans-rectal ultrasound (TRUS) was mandatory or allowed as an alternative to MRI in 55(59.8%) trials, of which 21 trials used TRUS only for local tumour staging. Only 43(46.7%) trials mandated MRI (Table 1) with 11(12.0%) having mesorectal fascia involvement as an eligibility criterion.
Treatment details and intensification regimens

Of the 92 trials, 8(8.7%) had a control arm that used a “standard” radiotherapy dose-fractionation and fluoropyrimidine regimen (Table 1). From a total of 104 experimental neoadjuvant treatment arms; 8 arms involved the addition of induction chemotherapy; 78 arms involved the intensification of concurrent chemoradiotherapy treatment only (by intensifying the radiotherapy dose and/or the addition of one or a combination of cytotoxic, biological or radiosensitising agent); and 18 arms involved both induction chemotherapy and concurrent chemoradiotherapy intensification.

Multiple dose and delivery schedules were used for concurrent chemotherapy intensification. For example, 34 treatment arms added oxaliplatin using 15 different schedules varying in dose, and number and frequency of cycles. Thirty-five experimental arms reduced the concurrent fluoropyrimidine dose (Table 1). Disregarding the differences in chemotherapy dose or delivery schedules, there were 29 different treatment agents/strategies tested involving intensification of induction chemotherapy, concurrent chemoradiotherapy or both.

Twenty-three different radiotherapy dose-fractionation schedules were used: eight standard fractionation schedules; six accelerated hyperfractionation; four hypofractionation; two simultaneous integrated boosts; and three brachytherapy or intra-operative radiotherapy boosts. Ninety-three (89.4%) experimental arms used a “standard” radiotherapy EQD2 equivalent dose, one arm a lower dose, and 10 arms a higher dose.
The modal recommended interval from the end of radiotherapy to surgery was 6-8 weeks (range: <1 – 23 weeks). Thirty-seven (40.2%) trials had a recommended minimum interval to surgery of under six weeks. Fifty-eight (63.0%) trials reported the actual interval to surgery, with an overall trial median of 6.9 weeks (range: 0.9 – 19.3 weeks).

**Statistical design and quality**

Seventy-five (81.5%) of the 92 trials were single-arm trials. Of the 17 multi-arm trials, 12 were randomised (Table 2), four trials consisted of sequential experimental arms and in one trial, treatment allocation was at “physician discretion”[93]. Only 51(55.4%) trials described their phase II statistical design allowing for sample size replication. Twenty-one (22.8%) trials did not describe how their sample size was derived.

Eighty-four (91.3%) trials defined a primary endpoint with 61(66.3%) trials using pCR rate. The remainder used 14 different primary endpoints broadly grouped into five themes: i] pathological response; ii] clinical response; iii] toxicity or treatment compliance; iv] R0 resection; and v] longer term outcomes (Table 2). The majority of trials (n=66, 71.7%) reported pathological regression grading with about one-third of trials using the Dworak tumour regression grading score[111].

All 92 trials reported pCR rates but only 25(27.2%) reported associated CIs or standard errors. Seventy-two (78.2%) trials provided their definition of pCR with 22 trials explicitly requiring local lymph node complete response. Longer-term endpoints (e.g. disease-free survival or overall survival) were reported in 41 trials.
The authors of 58 (63.0%) trials interpreted their results positively. Typical positive interpretations were that the experimental agent was “clinically active”, had “favourable” response rates, “safe” and “worthy of additional investigation”[50, 62, 75, 93]. Only 23 (25.0%) trials had a “positive” result meeting their original statistical hypothesis. Fifty-two (56.5%) trials had a “good” statistical design rating with the remainder failing to report their primary endpoint (N=8), and/or sufficient detail of their trial statistical design (N=38) or sample size calculations (N=20) to allow verification. Fifty trials (54.3%) had a “good” statistical compliance rating. Statistical non-compliance was due to authors’ positive interpretation of results despite not achieving pre-stated effect size (N=18), insufficient statistical design details to support authors’ conclusions (N=17), or non-compliance with the planned sample size (N=7). Only trials having both “good” statistical design and compliance to this design were rated “good” overall. Thirty-eight (41.3%) trials had a “good” overall statistical design and compliance (Table 2).

Meta-analysis
A total of 71 (77.2%) of the 92 trials were included in the meta-analysis (Appendix: Supplementary Table 2) after removing multi-arm trials (n=17) and those with no pCR events (n=4). Across single-arm trials of neoadjuvant long-course CRT treatment intensification, the pooled estimate of the pCR rate was 17.5% (95%CI: 15.7%-19.4%) (Figure 2). Higgin’s $I^2$ statistic (the proportion of total variance in pCR rates between trials due to heterogeneity) was substantial at 55.3%. Cochran’s Q-statistic, which tests the null hypothesis that the true treatment effect does not differ
between trials, was significant ($X^2=156.53$ (d.f.=70), $p<0.001$). Both tests indicate substantial between trial heterogeneity.

A multivariable meta-regression analysis explored the effect of the clinical design factors on the pooled estimate of the treatment effect, and between trial heterogeneity (Table 3). After fitting these covariates, Higgin’s $I^2$ statistic was reduced to 47.9% ($I^2_{\text{res}}$). The proportion of the between-trial variance explained by the covariates (adjusted $R^2$) was low at 15.6%. The multivariable meta-regression analysis (Table 3) suggests a significant association of higher radiotherapy EQD2 doses (not accounting for treatment duration) with a greater pCR rate (adjusted $p=0.025$). To assess the robustness of this result, the radiotherapy EQD2 was recalculated with a tumour repopulation dose increment to account for treatment duration and included in the model. In this model, radiotherapy dose was non-significant (adjusted $p=0.092$) and the adjusted $R^2$ value reduced to 6.82%. None of the other factors included in the model were shown to be significantly associated with pCR rate.

A sensitivity analysis further included 28 treatment intensification arms from 17 multi-arm trials. The pooled estimate of the pCR rate was concordant with the rate observed in the primary analysis (17.9%, 95%CI: 16.1%-19.6%); the estimate of the heterogeneity amongst reported effect sizes increased ($I^2=67.6\%$, $X^2=302.41$ (d.f.=98), $p<0.001$). The results of the meta-regression analysis were consistent with the primary analysis suggesting a significant association of higher radiotherapy EQD2 doses (not accounting for treatment duration) with a greater pCR rate (adjusted $p=0.041$).
A meta-analysis for long-term survival outcomes was not performed as only 42 of 92 trials provided data on such endpoints, and reporting of these outcomes was variable, e.g. loco-regional, disease-free, and/or overall survival were reported over seven different time-points (e.g. median survival, 1 to 6-year time-points).
Discussion

Our review highlights many problems in the conduct of treatment intensification phase II trials. The large number of single centre (57%) studies, the use of a single arm non-randomised design (82%) and the lack of trial registry registration (75%) are of particular concern.

Eligibility & staging - Over 90% of reviewed trials include patients with AJCC Stage II disease. These patients have a low local recurrence rate (<5%) but a high distant recurrence rate (~20%) with current treatment[3, 4, 112]. The incremental benefit of local treatment intensification only (70 of 92 trials reviewed) in this group is likely to be very small, and difficult to detect.

Although MRI staging was mandated in 46.7% of trials, only 12% included a threatened or involved mesorectal fascia in the definition of their eligible patient group. The MERCURY group have demonstrated the accuracy of MRI staging for assessing mesorectal fascia involvement and its prognostic value [113, 114].

Selection & reporting bias – Single centre single arm studies were commonly performed. Such studies are prone to significant selection bias and provide limited information to inform phase III study design. Single centre trials also tend to report larger treatment effects than multi-centre trials [115, 116], although this factor was not shown to be significant in our meta-regression analysis. Only 12 trials in this review were randomised, of which only eight trials used a standard chemoradiotherapy control arm.
In phase II trials, randomised trial designs allow comparison of novel treatment strategies for selection of the most promising for further testing[104], or the use of a “calibration” control arm[117] to minimise selection bias. Randomised trials also allow investigation for predictive biomarkers especially important in newer targeted agents. Clinical trials units are central to selecting and administering the most effective randomisation design.

A standard treatment control arm provides an important benchmark to establish true treatment differences. A 5% estimation error of the control effect size in a single-arm phase II trial could triple the false-positive error rate[118]. In trials involving combination treatments such as in rectal cancer, standard control arms prevent potential false impressions of benefit from improvements in standard care or better outcomes from selection biases [119, 120].

*Time to surgery* – There was a wide range of recommended intervals from the end of radiotherapy to surgery (range: <1-23 weeks) with limited evidence of the optimum interval. Current interest is whether a longer interval may produce further downstaging[121]. Our meta-regression analysis did not detect an association between the minimum recommended interval ≥6 vs. <6 weeks) and pCR rate, in keeping with the recent GRECCAR-6 trial finding no difference in pCR rates for 7 or 11 week intervals[122].

*Treatment intensification regimens* – From the multivariable analysis, the addition of a second concurrent chemotherapy agent did not show evidence of improved pCR rates. Reassuringly, compromising the concurrent fluoropyrimidine dose to facilitate
the addition of an experimental agent did not compromise pCR rates. However, the analysis suggests a significant association of higher radiotherapy EQD2 dose and higher pCR rates in keeping with radiation dose-response modelling[123]. This result should be interpreted with caution, in view of the large between-trial heterogeneity and the lack of validation of the radiobiological $\alpha/\beta$ and repopulation constants used, but may warrant investigation in future trials.

*Early efficacy endpoint* – It is clear that there is no validated surrogate end point for long-term cancer outcomes in rectal cancer. However, standardisation of efficacy measures are important and should be tailored to the clinical question and the patient sub-population studied. For example, in locally advanced rectal cancer where chemoradiotherapy is a peri-operative approach, pathological and imaging TRG may be more appropriate reporting than pCR. Pathological and MRI-defined TRG was prognostic for long term outcome in the MERCURY multicentre study[124, 125]. Other approaches being evaluated include the Neoadjuvant Rectal Score[126] and tumour cell density[127]. In contrast for an organ preservation trial, complete clinical response, colostomy and organ preservation rates should be reported[128].

For a single arm study, interpretation of pCR results is difficult with the lack of a contemporary standard treatment comparator, lack of an agreed pCR definition and the small sample size. Randomised studies would mitigate some of these problems. The European Rectal Cancer Consensus has also proposed a standardised pCR definition for future studies[129].
Trial design quality - Only 41.3% of the trials were rated as having “good” statistical design and compliance, with 44.6% and 22.9% of trials failing to report a statistical design and their sample size derivation respectively. Reporting of endpoints and their precision were variable with only 27.2% reporting CIs or standard errors. These are key elements in the reporting of clinical trial results, recommended by CONSORT[130]. A previous review of all phase II cancer trials found that positive findings were more frequently reported in trials with poor statistical design reporting[14].

Only 25.0% of the trials reported registration in a clinical trial registry. The FDA Amendments Act of 2007 mandates phase II trial registration on ClinicalTrials.gov[131, 132] to improve the disclosure of study design details, summary results reporting and minimise publication bias. A recent ASCO survey has demonstrated multiple common difficulties of maintaining clinical trial administrative and regulatory compliance throughout trial set-up, conduct and after completion[133]. Involvement of a clinical trial unit for all phases of phase II trial design, funding, registration, conduct and reporting would ensure the regulatory compliance, and reporting and statistical rigour of future trials.

Of the 29 different treatment intensification strategies, only five approaches - tested in 40 of the 92 phase II trials in this review - have led to subsequent testing in a phase III trial (published or registered on ClinicalTrials.gov (https://clinicaltrials.gov/)): i) concurrent oxaliplatin chemo-radiotherapy; ii) concurrent irinotecan chemo-radiotherapy; iii) radiotherapy dose-escalation via simultaneous integrated boost or hyperfractionated accelerated radiotherapy, and iv) induction FOLFOX or XELOX
chemotherapy. Only one of the six previously mentioned phase III trials testing the addition of concurrent oxaliplatin chemo-radiotherapy[7-12] has so far reported an improvement in disease free survival. Recent changes to NCCN guidelines regarding the use of neoadjuvant/induction chemotherapy are based only on phase II studies. The results of phase III trials are awaited.

This systematic review has several limitations. We restricted our review to published phase II trials thus excluding trials only presented in conference abstracts. We consider it appropriate to ensure prior peer-review of included trials. However, this may potentially exclude trials with negative results. Assessment of statistical quality was based purely on the published description thus may not be representative of the true trial design, while assessment of the author’s interpretation of their results was subjective. The trials evaluated in the meta-analysis were investigating varying forms of treatment intensification preventing assessment of any specific regimen. We did not evaluate the small number of phase II studies that used short course radiotherapy as we chose to evaluate intensification strategies associated with chemo-radiotherapy. Baseline tumour staging factors likely to be associated with pCR rates (e.g. TNM stage, mesorectal fascia involvement) could not be included in the meta-analysis as individual patient data was not available in 78 of 92 trials. In addition, as discussed, pCR has mixed utility limitations as an early efficacy measure and its definition varied between trials and compliance with the intention-to-treat principle in its reporting are unknown. Thus, the estimate of the pooled treatment effect, and subsequent meta-regression analysis should be interpreted with caution and the results serve for hypothesis generating purposes only. There are very likely
to be unknown sources of heterogeneity and biases unaccounted for within the analysis.

Possible solutions and recommendations

Our key trial design recommendations for future phase II trials in rectal cancer are:

- Careful consideration of eligibility criteria is recommended to ensure inclusion of patients most likely to benefit from the type of planned treatment intensification.
- MRI defined local staging inclusion criteria is strongly encouraged to improve patient eligibility, selection, and to allow treatment response assessment.
- Randomised multi-centre phase II trials with a “standard” treatment comparator control arm are strongly recommended. Many efficient phase II designs exist [119, 134-136].
- Appropriately statistically designed and sufficiently powered randomised phase II trials.
- Clinically relevant early efficacy endpoints should be used (e.g. TRG for locally advanced tumours; complete clinical response in organ preservation strategies).
- Clinical trial unit support is recommended for all stages of a phase II trial ensuring optimum trial design, statistical rigour, and regulatory compliance.

Our recommendations provide an opportunity to design a smaller number of higher quality multi-centre randomised phase II rectal cancer trials shortening the time-frame for this component of development and improving the likelihood of success of future phase III trials.
References


104. Wong SJ, Winter K, Meropol NJ et al. Radiation Therapy Oncology Group 0247: a randomized Phase II study of neoadjuvant capecitabine and irinotecan or capecitabine


Figure Captions

Figure 1: PRISMA flow diagram of literature search.

Figure 2: Forest plot of pCR rates and pooled estimate.
### Table 1: Summary of trial and treatment intensification details

<table>
<thead>
<tr>
<th>General</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
</tr>
<tr>
<td>Europe</td>
</tr>
<tr>
<td>Asia</td>
</tr>
<tr>
<td>North America</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient eligibility criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJCC stage</td>
</tr>
<tr>
<td>Stage II/III</td>
</tr>
<tr>
<td>Stage III only</td>
</tr>
<tr>
<td>Not stated</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Local staging modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI mandatory</td>
</tr>
<tr>
<td>43</td>
</tr>
<tr>
<td>MRI optional</td>
</tr>
<tr>
<td>No MRI</td>
</tr>
<tr>
<td>Not stated</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mesorectal fascia involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threatened, Involved or 'Unresectable'</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>Not assessed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control arm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard treatment control arm</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment intensification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addition of induction chemotherapy</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Cytotoxics only</td>
</tr>
<tr>
<td>Biological agent only</td>
</tr>
<tr>
<td>Combination</td>
</tr>
<tr>
<td>Chemoradiotherapy intensification</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Addition of cytotoxic</td>
</tr>
<tr>
<td>Addition of biologic</td>
</tr>
<tr>
<td>Addition of radiosensitiser</td>
</tr>
<tr>
<td>Radiotherapy dose-intensification</td>
</tr>
<tr>
<td>Combination of above</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other treatments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjuvant chemotherapy</td>
</tr>
<tr>
<td>Mandatory</td>
</tr>
<tr>
<td>Optional</td>
</tr>
<tr>
<td>No recommendation</td>
</tr>
<tr>
<td>Standard treatment dose compromised</td>
</tr>
<tr>
<td>No compromise in dose</td>
</tr>
<tr>
<td>Fluoropyrimidine compromised</td>
</tr>
<tr>
<td>Table 2: Trial statistical details and design</td>
</tr>
<tr>
<td>---------------------------------------------</td>
</tr>
<tr>
<td><strong>No of studies</strong></td>
</tr>
<tr>
<td>No of treatment arms</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>Randomised</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Statistical design</td>
</tr>
<tr>
<td>Described</td>
</tr>
<tr>
<td>Simon 2-stage</td>
</tr>
<tr>
<td>Fleming 1-stage</td>
</tr>
<tr>
<td>Gehan 2-stage</td>
</tr>
<tr>
<td>Other(^a)</td>
</tr>
<tr>
<td>Not referenced but sample size reproducible</td>
</tr>
<tr>
<td>Not described</td>
</tr>
<tr>
<td>Sample size calculation</td>
</tr>
<tr>
<td>Stated</td>
</tr>
<tr>
<td>Not stated</td>
</tr>
<tr>
<td>Primary endpoint</td>
</tr>
<tr>
<td>Stated</td>
</tr>
<tr>
<td>pCR</td>
</tr>
<tr>
<td>Pathological response/ downstaging</td>
</tr>
<tr>
<td>Toxicity/ Treatment compliance</td>
</tr>
<tr>
<td>LRFS/ DFS/ Local control</td>
</tr>
<tr>
<td>R0 resection</td>
</tr>
<tr>
<td>Clinical response/ downstaging</td>
</tr>
<tr>
<td>Not stated</td>
</tr>
<tr>
<td>Regression grading system used</td>
</tr>
<tr>
<td>Not used</td>
</tr>
<tr>
<td>Dworak</td>
</tr>
<tr>
<td>TNM downstaging</td>
</tr>
<tr>
<td>Mandard</td>
</tr>
<tr>
<td>JSCCR</td>
</tr>
<tr>
<td>AJCC</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Author interpretation</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Neutral</td>
</tr>
<tr>
<td>Negative</td>
</tr>
<tr>
<td>Overall statistical design and compliance quality</td>
</tr>
<tr>
<td>Good</td>
</tr>
<tr>
<td>Poor</td>
</tr>
</tbody>
</table>

\(^a\)Other designs used in only one study each (see Supplementary Table 1)
Table 3. Multivariable meta-regression analysis for pooled pCR rate

| Covariate                                                                 | Coefficient | SE  | t      | p>|t| | Adjusted p** | 95% CI          |
|---------------------------------------------------------------------------|-------------|-----|--------|-----|---------------|----------------|
| Standard fluoropyrimidine dose compromised (Yes vs. No)                    | -0.011      | 0.021 | -0.54  | 0.592 | 0.993         | -0.053, 0.031  |
| Recommended minimum interval from the end of (chemo)radiotherapy to surgery (≥ 6 vs. < 6 weeks) | 0.037       | 0.022 | 1.72   | 0.091 | 0.389         | -0.006, 0.081  |
| Radiotherapy dose*                                                         | 0.008       | 0.003 | 2.91   | 0.005 | 0.025         | 0.002, 0.013   |
| Addition of a second concurrent chemotherapy agent (Yes vs. No)           | 0.029       | 0.026 | 1.10   | 0.278 | 0.830         | -0.024, 0.081  |
| Addition of neoadjuvant chemotherapy before or after (chemo)radiotherapy (Yes vs. No) | 0.001       | 0.023 | 0.06   | 0.956 | 1.000         | -0.045, 0.047  |
| Single vs. multi-centre studies                                           | -0.002      | 0.020 | -0.08  | 0.938 | 1.000         | -0.041, 0.038  |

*Based on EQD2; ** Permutations = 20000

$I_{res}^2 = 47.9\%$; Adjusted $R^2 = 15.6\%$; Model $F (6, 58) = 2.08$; Prob $> F = 0.07$
926 articles identified through database search

→ 298 duplicates removed

→ 628 articles after duplicates removed

→ 516 articles excluded based on irrelevant title, abstract or citation:
- 2 excluded involving short-course radiotherapy
- 24 excluded being a Phase I/II trial

→ 112 articles included for full text analysis

→ 5 additional articles identified from manual review of reference sections

→ 117 articles included for full text analysis

→ 92 articles included for data extraction and formal review

→ 5 excluded being a Phase I/II trial

→ 1 excluded based on being a non-rectal cancer study

→ 12 excluded based on duplicate populations

→ 5 excluded based on no radiotherapy or radiotherapy EQD2 <40Gy

→ 2 excluded being a clinical trial note with no results
**NOTE:** Weights are from random effects analysis