Alternative clinical end points in rectal cancer—are we getting closer?

R. Glynne-Jones†*, S. Mawdsley, T. Pearce & M. Buyse

†Centre for Cancer Treatment, Mount Vernon Hospital, Northwood, Middlesex, UK; ‡Gray Laboratory Cancer Research Trust; †PDL Bio Pharma; ‡International Drug Development Institute (IDDI), Inc

Received 8 May 2006; revised 8 June 2006; accepted 14 June 2006

Background: In rectal cancer a high risk of local recurrence has been reported for patients treated by surgery alone. It is also recognised that 20%–40% of rectal cancer patients continue to develop distant metastases and die, even when a very low risk of local recurrence has been achieved with the use of preoperative radiotherapy and total mesorectal excision (TME). Hence, the current design of randomised trials in rectal cancer continues to use the standard end points of local control and survival. This strategy is time-consuming. The recently published EORTC 22921 trial, which compared radiotherapy with chemoradiotherapy and tested the role of postoperative adjuvant chemotherapy, has taken 14 years from planning to results. The aim of this review was to use the evidence from both phase II and phase III trials to provide a comprehensive survey of alternative clinical trial end points in rectal cancer, where preoperative chemoradiation has now become the standard treatment. We describe their strengths and weaknesses. Some are clearly defined, easy to assess and can be obtained early, because surgical resection usually takes place within 6–8 weeks of the completion of treatment. Some pathological response end points reflect biological activity, although their effect on survival has yet to be validated in randomised controlled trials. We will propose measurement and analytical techniques for minimising bias and intra- and inter-observer variability of the non-validated end points in the hope of basing these judgements on as firm a ground as possible.

Methods: A literature search identified both randomised and non-randomised trials of preoperative radiation therapy (RT) and chemoradiation therapy (CRT) in rectal cancer from 1993 to 2005. The aim was to find those studies that documented potential alternative end points.

Results: Pathological parameters have been used as early end points to compare studies of preoperative radiotherapy or chemoradiation. In the light of the German CAO/ARO/AIO-94 study, which demonstrated an improved therapeutic ratio for preoperative treatment, enthusiasm for preoperative chemoradiation in the management of rectal cancer is increasing. Current evidence cannot indicate whether the degree of response to chemoradiation (e.g. complete pathological response; downsizing the primary tumour; sterilizing the regional nodes; tumour regression grades or residual cell density) or the achievement of a curative resection (CRM/R0 resection) is the best early clinical end point. Problems with these end points include lack of structured measurement and analysis techniques to control for intra- and inter-observer variation and lack of validation as surrogates for long-term clinical end points such as local control and survival. However, retrospective studies in rectal cancer have confirmed a strong association between the presence of microscopic tumour cells within 1 mm of the CRM and increased risks of both local recurrence and distant metastases. Further end points of current clinical relevance for which adequate methodologies for assessment are lacking include sphincter sparing end points, and assessment of long-term toxicities, ano-rectal function and their specific impact on quality of life. Recommendations are made as to the most appropriate information, which should be documented in future trials.

Conclusions: Pathological complete response following preoperative chemoradiation does not reliably predict late outcome. There are other events not mediated through this end point and there are also unintended effects (often an excess of non-cancer related deaths). Disease-free survival currently remains the best (because it is relatively quick) primary end point in designing randomised phase III studies of preoperative chemoradiation in rectal cancer, although it is necessary to control for postoperative adjuvant chemotherapy. However, the CRM status can substantially account for effects on disease-free and overall survival after chemoradiation, radiation or surgery alone. Hopefully, randomised controlled trials, which utilise these alternative clinical end points, will in future determine the precise percentages of the effect of different chemoradiation schedules on disease-free and overall survival.
Key words: chemoradiation, circumferential resection margin, complete pathological response, radiation, rectal neoplasms, surrogate endpoints

Introduction

The effective chemotherapeutic armamentarium in the treatment of colorectal cancer has increased in the past decade from one option i.e. 5-fluorouracil (5-FU) to five with the recent availability of irinotecan, oxaliplatin, cetuximab and bevacizumab. Patients with metastatic disease potentially receive two to four lines of therapy, which has extended median survival from approximately 12 to 24 months [1–4]. These regimens are now being extrapolated into the adjuvant setting.

There has also been a recent paradigm shift in rectal cancer from postoperative to preoperative treatment as a result of a number of studies culminating in the German CAO/ARO/AIO-94 protocol [5]. In this trial preoperative chemoradiation has been convincingly shown to be more effective in achieving local control, with less acute and late toxicity than postoperative chemoradiation.

Surgical technique for rectal cancer has also improved with the use of total mesorectal excision (TME). TME can reduce local recurrence rates to below 10% and, with the addition of radiotherapy, to less than 3% [6–8], even in stage III patients [9]. Yet in trials 10%–20% of patients with resectable disease (as defined by clinical staging) still fail to achieve a curative (histologically confirmed) R0 resection in terms of the circumferential margin [10–12].

Locoregional relapse has been the predominant form of treatment failure; but now, with improved local surgical and radiotherapeutic techniques, locoregional relapse is exceeded by the rate of development of systemic metastases [5, 13, 14]. However, if long-term local control can be achieved in more than 90% of patients, then designing trials to distinguish therapies for this end point will require large sample sizes and enormous resources. Locoregional control has thus been questioned as the most relevant primary end point for resectable rectal cancer [15].

Prolonging overall survival remains the ultimate goal of cancer therapies. In clinical trials, however, survival may also be a problematic end point because the current availability of multiple effective lines of systemic therapy obscures the impact of study treatment upon survival. Furthermore, the interval from the end of recruitment to primary efficacy analysis is protracted, such that subsequent trials taking the best therapy from the previous trial cannot begin until years after the previous trial has completed recruitment (the same problem exists for disease-free survival).

As a case in point, the European Organization for Research and Treatment of Cancer (EORTC) Phase III study 22921 randomised 1011 patients with T3/T4 resectable rectal cancer between long fractionation preoperative radiotherapy with or without concurrent 5-FU and leucovorin. This trial was designed in 1991, but completed accrual in April 2003. The end points were overall and disease-free survival and local control. The toxicity results and early end points of response are available [16, 17], but the primary efficacy end points have only recently been presented in abstract [14]. Fourteen years is a long time to wait in oncology!

In contrast, neo-adjuvant therapies offer the option for using alternative early end points of response based on histopathological parameters. These end points include the achievement of a curative (R0) resection, a negative circumferential resection margin (CRM), the degree of responsiveness to therapy, e.g. pathological complete response or downsizing, or tumour regression grading and early assessment of toxicity. The potential advantages include being available in the short term, because surgical resection usually takes place within 6–8 weeks of the completion of treatment, and being objectively measurable. At least one surgical trial has used short-term end points [10]. Current problems include lack of structured measurement and analysis techniques to control for intra- and inter-observer variation, and lack of validation as surrogates for long-term clinical end points such as local control and survival. Further end points of current clinical relevance for which adequate methodologies for assessment are lacking include sphincter sparing end points, and assessment of long-term toxicities, anorectal function and their specific impact on quality of life. To date, these alternative end points have, to a large extent, remained post hoc hypotheses and have not been tested in the randomised preoperative setting.

The aim of the study was to provide a comprehensive survey of end points in trials for rectal cancer investigating preoperative treatment strategies, and to describe the strengths and weaknesses of each. These end points can be reviewed and confirmed by blinded experts and reflect biological activity, although their effect on survival has yet to be validated in randomised controlled trials. These end points will not be influenced by the potential variety of postoperative treatments. We propose measurement and analytical techniques for minimizing bias and intra- and inter-observer variability of the non-validated end points.

New phase III trials will have to select a control arm from previous phase III trials using surrogate end points because the definitive end points will not be available in a relevant timeframe. The subsequent trial, therefore, must include as a prospective objective the validation of the surrogate end point used to select the control arm. We will lay out a general approach to such validation.

Selection of trials

A literature search was performed using Medline and Cancerlit from 1993 to 2005 inclusive, supplemented by hand searching of abstracts from the Proceedings of the American Society of Clinical Oncology and other international meetings (ASTRO, ESTRO, ESMO and ECCO), and cross-referencing to provide evidence for this discussion from randomised and non-randomised trials of preoperative radiation therapy (RT) and chemoradiation therapy (CRT) in rectal cancer. Based on the titles of the articles and excluding those describing recurrent cancer, local excision, follow-up strategies and prostate cancer, we retained articles and abstracts in English, German and French. We read the full text of 143 articles, which seemed likely from the title to offer information regarding the defined...
end points of overall survival, disease-free survival, pathological complete response, circumferential resection margin status, R0 resection status, surgical morbidity, functional outcome and quality of life. If trial results were subsequently updated, information from the more recent publication was used. Small trials treating less than 20 patients were excluded. Data was checked for accuracy by the authors but not by independent persons.

**survival**

Trial end points based on overall survival which require waiting a minimum of 7–10 years to observe reliable results, are inevitably lengthy and expensive. In trials of chemotherapy for patients with metastatic colorectal cancer, it has become difficult to demonstrate the individual benefit of any new drug. The effects of crossover and the possibility of confounding the survival end point from sequential lines of effective systemic treatment, make it a gruelling task to identify what is the particular relevant change from which a small benefit in survival is observed.

An increasingly popular end point is disease-free survival (DFS). DFS is increasingly chosen as the end point for the US Food and Drug Administration and NCI trials of adjuvant chemotherapy in colon cancer. DFS is an appropriate end point, not only because DFS correlates with overall survival (see below) but also because preventing local recurrence is worthwhile in itself. DFS correlates with overall survival (coefficient of correlation >0.90) and in a recently reported meta-analysis including some 13 000 patients with resectable colorectal cancer, the benefits of various experimental therapies on 3-year DFS were shown to correlate with the benefits on 5-year survival (coefficient of correlation >0.90) [18]. These two tight correlations suggest that 3-year DFS can be considered a valid surrogate for 5-year survival, at least for the class of agents included in the meta-analysis [19]. DFS could be used in a preoperative chemoradiation study, although there may be difficulties with defining local recurrence (see below). Most postoperative adjuvant studies in colorectal cancer now use the end point of 3-year DFS, as 90% of the recurrences will have occurred within 3 years. However, preoperative chemoradiation treatment may serve to delay rather than abolish local recurrence. Also, outcome will probably relate to both the intensity of preoperative chemoradiation, as well as the use of more effective postoperative adjuvant chemotherapy.

**local control**

Local control has in the past been considered the ‘gold standard’ end point in rectal cancer trials. Yet the diagnosis of local recurrence may prove difficult even if regular pelvic CT scans are mandated within a trial, because substantial soft tissue masses are often seen in the posterior pelvis following preoperative chemoradiation and surgery. There are also numerous methodological differences in defining and reporting local recurrence [20]. Median follow-up after chemoradiation will also have to be maintained over a much longer period than many trials have yet reported [21].

The Dutch CKVO 95-04 trial study reported a local recurrence rate of 2.4% at 2 years [8] and 5.8% at 5 years [22] with a combination of preoperative radiotherapy and TME. The German CAO/ARO/AIO-94 study [5] demonstrates a local recurrence rate at 5 years of only 6% after preoperative chemoradiation. Other recently reported trials of chemoradiation also showed low recurrence rates of 6%–8% at 5 years [14, 23], which were significantly improved with radiotherapy alone. If such low rates can be widely achieved in trials, it seems unlikely that this end point will remain a useful discriminatory tool for comparing treatment strategies, at least in resectable cancer.

The NSABP R-04 phase III randomised study in T3/T4 resectable rectal cancer compares preoperative continuous infusional 5-FU with capecitabine, given concurrently with radiation, and is integrating oxaliplatin into this schedule. This study is planned to recruit several thousand patients defined as resectable by trans-rectal ultrasound. The trial has recently moved from using loco-regional recurrence as a primary end point to a combination end point of pathological complete response and disease-free survival.

**distant control**

Although the MRC2 study [79] showed a significant reduction in metastatic disease for the 279 patients who received preoperative radiotherapy, this advantage was not observed in the 3029 patients treated in the Swedish and Dutch trials (see Table 1). There is no evidence that 5-FU-based preoperative chemoradiation regimens achieve any effect on metastatic disease, because the radiotherapy versus chemoradiotherapy trials have reported similar rates of distant metastases and disease-free survival [14, 23]. Perhaps this finding reflects the fact that the chemoradiation regimens investigated to date incorporate low-dose intensities of 5-FU alone (and not in combination) and low total doses of systemic therapy.

**the extent of resection/R0 resection**

The presence of residual tumour following surgical resection strongly determines future outcome. The American Joint Committee on Cancer Prognostic Factors Consensus Conference defines R0 as the absence of residual disease, R1 as residual microscopic disease and R2 as residual macroscopic disease. The UICC definition of a curative resection is similar in that it demands microscopically negative resection margins (UICC staging on cancer). A CRM of 0 mm would, therefore, categorise a resection as R1.

**the radial or circumferential resection margin**

Traditionally histopathologists reported only the proximal and distal bowel margins in rectal cancer as these were the parameters surgery focused on. With TME and its goal of containing all tumours within an enclosed fascial compartment, Quirke defined the risks of local recurrence in rectal cancer by carefully measuring the minimum distance between the
nearest extent of the tumour in the resected specimen and the CRM [24 25]. Now that TME is routinely practised, reporting on the CRM is an essential component of surgical quality control.

A minimum distance of 1 mm appeared to discriminate between a high (85%) or low risk (3%) of local recurrence. Further evidence from this group [26] lent support to this hypothesis. Updating confirms the importance of the CRM despite the introduction of TME techniques [27].

The Dutch study questions whether a 2 mm margin might be a more appropriate cut-off point [28]. However, the numbers of patients to support this view are very small (47 versus 53 patients, with and without radiotherapy, respectively) and not statistically significant. Only one other study suggests the relevance of a 2 mm margin [29].

Failure to achieve a negative CRM not only confers a high risk of local recurrence, but also doubles the risk of developing metastases; see Tables 1 and 2 [8, 30, 31]. Other studies suggest that patients with an involved CRM often develop and die of metastatic disease before local recurrence is observed [30, 32]. Furthermore, data from phase II chemoradiation studies show that patients who fail to achieve a CRM of >1 mm after chemoradiation virtually all relapse, with a DFS of only 10% at 3 years [33] and a high risk of metastases [34].

It is not always clear, however, whether a positive CRM may reflect inadequate surgery or more advanced aggressive disease. A recent study by the Association of Coloproctology of Great Britain and Ireland, involving 1036 patients from 37 centres undergoing rectal cancer resection, reported CRM involvement rates of 7.5%, 16.7% and 31.7% following anterior resection, abdomino-perineal excision of rectum (APER) and Hartmann’s procedure, respectively [35].

Not all histopathologists have adopted these meticulous techniques [36]. A positive CRM is probably seriously under-reported, unless the pathologist employs multiple slice analysis [37]. A study examining the degree of variability within the North Central Cancer Treatment Group found the extent of radial spread was poorly documented [38]. The College of American Pathologists Consensus Statement acknowledges that measurement of the CRM is not routinely employed in the USA [39]. Yet, of all risk factors for local recurrence and metastatic disease, the CRM could potentially be influenced by

### Table 1.

<table>
<thead>
<tr>
<th>Phase II/observational studies</th>
<th>No. of patients</th>
<th>CRM --/CRM + (%)</th>
<th>Local recurrence (%)</th>
<th>Metastases (%)</th>
<th>3-year DFS (%)</th>
<th>OS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quirke [24] (non TME)</td>
<td>52</td>
<td>--ve 73</td>
<td>3</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 27</td>
<td>86</td>
<td>No data</td>
<td>No data</td>
<td></td>
</tr>
<tr>
<td>Adam [26]</td>
<td>141</td>
<td>--ve 75</td>
<td>10</td>
<td>No data</td>
<td>No data</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 25</td>
<td>78</td>
<td>No data</td>
<td>No data</td>
<td>24</td>
</tr>
<tr>
<td>De Haas-Kock [40]</td>
<td>253</td>
<td>--ve 88</td>
<td>8</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 12</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hall [30] (all had TME)</td>
<td>152</td>
<td>--ve 87</td>
<td>11</td>
<td>17</td>
<td>74</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 13</td>
<td>15</td>
<td>35</td>
<td>41</td>
<td>46 at 4 years</td>
</tr>
<tr>
<td>Wibe [31] (TME)</td>
<td>686</td>
<td>--ve 90</td>
<td>5</td>
<td>12</td>
<td>No data</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 10</td>
<td>22</td>
<td>40</td>
<td>No data</td>
<td>66 at 3 years</td>
</tr>
<tr>
<td>Birbeck [27]</td>
<td>586</td>
<td>--ve 72</td>
<td>10</td>
<td>No data</td>
<td>No data</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 28</td>
<td>38</td>
<td></td>
<td></td>
<td>29</td>
</tr>
</tbody>
</table>

### Table 2.

<table>
<thead>
<tr>
<th>Randomised phase III studies</th>
<th>No. of patients</th>
<th>CRM --/CRM + (%)</th>
<th>Local recurrence (%)</th>
<th>Metastases (%)</th>
<th>3-year DFS (%)</th>
<th>OS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classic [10]</td>
<td>794 (300 rectal)</td>
<td>--ve 85</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 15</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>CR07</td>
<td>1300+</td>
<td>--ve 87</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 13</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>Dutch CKVO 95–04 11</td>
<td>1861</td>
<td>--ve 82</td>
<td>8</td>
<td>16</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ve 18</td>
<td>3 at 2 years</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>Polish Study; Bujko [12]</td>
<td>RT alone 155</td>
<td>--ve 87</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td></td>
<td>CRT 157</td>
<td>--ve 96</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
</tbody>
</table>
the overall impact and quality of treatment (surgery, chemotherapy and radiation).

Results from the first published randomised comparison of preoperative short-course RT versus conventional long-fractionation CRT suggest an advantage to the latter approach: positive CRM rates were 13% for short fractionation RT versus 4% for long fractionation CRT [12].

The CRM is the easiest, most reliable and reproducible early end point that predicts both local recurrence and metastatic disease. The authors believe that the best international definition of R0 should be an assessment of distal and radial margin of ≥1 mm, as determined using standardised multiple slice analysis by a blinded central pathologist. This emphasis on the circumferential margin should be mandated in future studies in rectal cancer. The distance between the primary tumour and the radial margin should be documented, e.g. a histologically involved margin, 0–1 mm, 1–2 mm and more than 2 mm. Formal attempts to validate R0/CRM negative resection as a surrogate end point for disease recurrence will be possible when data are available from enough randomised studies to perform a meta-analysis.

**quantification of the degree of response**

Several parameters have been considered to quantify tumour regression after preoperative chemoradiation, which include pathological complete response, T and N stage downstaging, residual cell density and histological regression grades.

**complete pathological response**

Complete pathological response (CPR) is an attractive end point, but does CPR translate into improved local recurrence rates, or prolonged disease-free or overall survival when compared with partial responders? Phase III studies such as the German CAO/ARO/AIO-94 study [41] and the NSABP R03 study [42] have evaluated the achievement of a complete pathological response (CPR) following preoperative chemoradiotherapy (CRT) and surgery for locally advanced rectal cancer. These studies suggest that a better outcome in terms of overall survival and disease-free survival is achieved for patients who achieve a complete pathological response after chemoradiation. Some phase II trials seem to suggest that low dose-intensity systemic chemotherapy combined with purely locally-acting radiation therapy implies the elimination of distant microscopic disease [43–46].

However, in randomised trials despite the observation of higher CPR rates, the addition of chemotherapy has had no impact on overall or disease-free survival [14, 23]. Both the French EORTC 22921 study [14] and the FFCD 9203 study [23] suggest that the addition of 5-FU to radiation increases the CPR rate from 3% to 13.7% and from 3% to 11.7% in resectable T3/T4 rectal cancer, respectively. In these studies chemoradiation led to a reduction in local recurrence but did not impact on disease-free or overall survival.

We may be overstaging and overtreating early tumours, which have an excellent prognosis without any adjuvant treatment. Three recent randomised trials had eligibility that required resectable T3/T4 tumours, and contained a control arm that represented either initial surgery, or short fractionation preoperative radiotherapy, which does not cause down-staging. These trials appear to have over-staged a significant proportion of patients. The proportion of T1/T2 tumours was 18% in the German study [5] and 39% in the Polish study [12]. Even if these smaller, earlier tumours were more likely to achieve a CPR, it may not be relevant as T1/T2 N0 tumours would be expected to have a good prognosis anyway.

An update of a phase I/II study with irinotecan [47] suggests that the CPR rate may be much higher for small tumours <5 cm. They showed that 11/28 (39%) of small tumours achieved CPR, compared with only five of 30 (17%) tumours larger than 5 cm. It could, therefore, be argued that inaccurate clinical staging contributes to high CPR rates.

It remains unclear whether we are actually altering the natural history of the disease with preoperative chemoradiation or just revealing a biological subgroup that would have had a very favourable outcome anyway. The available data suggests that modest increases in PCR do not impact on survival. Only data on long-term survival for this group of patients in the context of a randomised trial or from a meta-analysis will provide this information.

Table 3 shows CPR rates from the randomised trials in rectal cancer. These trials have either compared chemoradiation with

<table>
<thead>
<tr>
<th>Trial</th>
<th>RT dose(Gy)</th>
<th>No.</th>
<th>Chemoradiation</th>
<th>Radiation</th>
<th>Surgery alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>EORTC 1984 [48]</td>
<td>34.5</td>
<td>247</td>
<td>5% (6/126)</td>
<td>2.5% (3/121)*</td>
<td>—</td>
</tr>
<tr>
<td>INT 0147</td>
<td>50.4</td>
<td>53</td>
<td>Closed early</td>
<td>—</td>
<td>No data</td>
</tr>
<tr>
<td>NSABP R03 [49]</td>
<td>45</td>
<td>267</td>
<td>16% (21/130)</td>
<td>—</td>
<td>0/137</td>
</tr>
<tr>
<td>Frykholm [50]</td>
<td>40 (split)</td>
<td>70</td>
<td>12% (3/29)</td>
<td>4% (1/27)*</td>
<td>—</td>
</tr>
<tr>
<td>CAO/ARO/AIO-94 [5]</td>
<td>50.4</td>
<td>823</td>
<td>8% (33/415)</td>
<td>—</td>
<td>0%</td>
</tr>
<tr>
<td>Polish Trial [11]</td>
<td>50</td>
<td>316</td>
<td>16% (22/138)</td>
<td>1% (1/138)</td>
<td>—</td>
</tr>
<tr>
<td>EORTC 22921 [14]</td>
<td>45</td>
<td>1011</td>
<td>13.7% (65/473)</td>
<td>5.3% (25/476)</td>
<td>—</td>
</tr>
<tr>
<td>FFCD 9203 [12]</td>
<td>45</td>
<td>762</td>
<td>11.7% (40/370)</td>
<td>3.7% (13/363)</td>
<td>—</td>
</tr>
<tr>
<td>Berrocal [52]</td>
<td>45–50</td>
<td>155</td>
<td>13.2% (10/76)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Not significant.
Factors associated with a greater chance of complete pathological response not directly related to therapy include initial T stage overall tumour size, the interval between the completion of radiotherapy and surgery [51] and the meticulousness of the pathological assessment. A recent review of phase II and III trials of preoperative chemoradiotherapy [53] showed that on multivariate analysis only the mode of delivery of 5-FU (infusion rather than bolus), the use of a second drug in addition to 5-FU and the total radiation dose were associated with a higher rate of pathological complete response.

The Lyon Study elegantly demonstrated that a longer interval (6–8 weeks versus 2 weeks) increases the CPR rate in tumours without reducing the local recurrence rate or survival [54, 55]. Further extension beyond 8 weeks did not offer further advantages in downstaging. In contrast, a study using irinotecan in a preoperative chemoradiation schedule did not find a statistically significant advantage in terms of downstaging for 10–14 weeks compared with 4–8 weeks [56]. A further study from France also suggested that the interval between radiotherapy and surgery, although allowing greater potential for sphincter sparing surgery, may be detrimental to survival if extended too long [57]. More than 15 weeks between diagnosis and surgery, and an interval longer than 6 weeks between the completion of radiotherapy and surgery may reduce overall survival. The optimal interval remains controversial.

A CPR means that the pathologist cannot demonstrate any viable cancer cells within the operative specimen. However, the pathologist’s success or failure to find viable cells, depends on how assiduous is their search and how many sections are processed. Currently, a European multicentre chemoradiation study (CORE study) is attempting to define, prospectively, a standard procedure for defining CPR.

Even if cells remain, we have no accurate way of assessing their potential for division or their true viability. Lakes of mucin can be seen in the specimen either within a fibrotic scar following chemoradiation or completely replacing lymph nodes. If the mucin is not associated with adjacent malignant cells, we have no accurate way of assessing their potential for division or their true viability. Lakes of mucin can be seen in the specimen either within a fibrotic scar following chemoradiation or completely replacing lymph nodes. If the mucin is not associated with adjacent malignant cells, should pathologists have classified these changes as ypT3 N1 (Dukes C1) or as a CPR?

**residual cell density (RCD) and tumour regression grade (TRG)**

The histopathological appearances following preoperative chemoradiation in rectal cancer have been poorly documented and a standardised approach to characterising these changes is lacking. This difficulty has been addressed in a recent multicentre phase II study—CORE [58]. The majority of patients achieve less than a PCR and there is a wide variety of definitions and techniques for identifying and scoring the residual cancer. A tumour regression scale described in esophageal cancers following chemoradiation used a five-point scale basing the grades on the amount of residual tumour cells and the extent of fibrosis in the bowel wall achieved [59]. This system scored the degree of cytological changes (nuclear pyknosis or necrosis), the stromal changes (fibrosis), the observation of inflammatory and giant cell granuloma. In addition, some authors have categorised the finding of only a few microscopic foci of cancer outside the muscularis propria as Tmic rather than T3 [54, 60, 61], often only in a single block. Some authors have suggested that the presence of residual microscopic focci compared with gross residual tumour confers a better prognosis [62]. Other small studies have attempted to test the inter-observer reliability of this scoring system [63], which appeared high.

The most favourable regression grade represents complete sterilisation of the tumour or CPR, which is discussed above, but the regression grade system also captures the remaining 85% of patients.

Variations of the Mandard regression grading with three or four grades have also been employed [61, 62, 64] and claimed to be more easily reproducible. Sadly, TRG has no internationally accepted system of classification. However, if the TRG numbers are ignored, then the regression grades can be compared in terms of complete response, good response, moderate response, poor response and no response. Other investigators have measured the percentage of residual tumour in the specimen and lymph nodes [65] or have tried to grade tumour necrosis as a percentage of the tumour mass. Bouzourene [66] demonstrated that patients who exhibited large amounts of tumour necrosis appeared to survive longer, although the improvement did not reach statistical significance. On multivariate analysis tumour regression grade remained an independent prognostic indicator for long-term local tumour control.

A regression grading system used by Rodel [67] and the German CAO/ARO/AIO-94 protocol [68, 69] demonstrate that the degree of regression at 6 weeks influences subsequent outcome. In this exploratory and initially unplanned analysis, complete (TRG 4) and intermediate pathologic response (TRG 2 + 3) were associated with an improved DFS after preoperative CRT. The authors recommend that TRG assessment should be implemented in the standard pathological format of evaluation and prospectively validated in further

<table>
<thead>
<tr>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence of tumour—residual fibrosis</td>
<td>Rare tumour cells scattered throughout fibrosis</td>
<td>Increase in number of residual cells—predominant pattern of fibrosis</td>
<td>Residual tumour out growing fibrosis</td>
<td>No evidence of tumour regression</td>
</tr>
<tr>
<td>Complete response</td>
<td>Good response</td>
<td>Moderate response</td>
<td>Poor response</td>
<td>No response</td>
</tr>
</tbody>
</table>

Table 4. Mandard scale of regression in oesophageal cancer
studies. However, it remains unclear how reproducible RCD or TRG is as an end point, outside the category of CPR, and whether RCD is a prognostic factor or not [70, 71].

**clinical/radiological downstaging**

Downstaging a tumour after chemoradiation implies that accurate and reproducible methods of preoperative and postoperative staging are available. Many phase II trials describe high rates of downstaging. Are they justified? Endoscopic rectal ultrasound (ERUS) can provide some imaging of the para-rectal lymph nodes, but obturator nodes and lateral pelvic wall nodes are often difficult to see and assess either with RUS or spiral CT imaging. Finally, MRI may accurately quantify the proximity of a tumour to the mesorectal envelope and, therefore, predict the likelihood of a R0 resection [72, 73]. So there are inherent methodological problems of comparing preoperative imaging with post-operative pathological data.

Is downstaging the T stage as important as downstaging the N stage? Does the extent of downstaging predict for local control or overall survival? The prognostic significance of post chemoradiation stage following preoperative chemotherapy and radiation for advanced or recurrent rectal cancers has recently been extensively explored in a retrospective study [43].

**sterilising lymph nodes**

In patients with rectal cancer treated with surgery alone, the status of the lymph nodes remains the most important factor in predicting future outcome. Both lymph node invasion, and the number of lymph nodes involved are independent prognostic factors for survival [74–76]. When four or more lymph nodes are involved, the prognosis is much worse. In addition, the overall number of lymph nodes reported is of prognostic significance in rectal cancer [77] whether involved or not. It has been suggested that the higher the number of lymph nodes examined, the more accurate the staging.

Randomised studies of preoperative radiotherapy versus surgery alone [78, 79] and preoperative chemoradiation versus preoperative radiation [48] have demonstrated significant reductions in the proportion of patients with Dukes C histology. In the German CAO/ARO/AIO-94 study, after surgery there was a lower percentage of tumours with positive lymph nodes in the preoperative chemoradiation arm compared with those who proceeded straight to surgery: 29% versus 44%, respectively [80].

Following preoperative chemoradiotherapy, pathologists often cannot find sufficient lymph nodes to pronouce on stage. Retrospective studies have shown lymph nodes are smaller and are less likely to contain malignant cells [81]. Pathologists also report mucin lakes in lymph nodes following chemoradiation and occasionally lymph nodes replaced by fibrosis. In summary, the level of downstaging or lack of response—or, more precisely, failure to sterilise lymph nodes—may provide a clinically useful surrogate end point for future studies of chemoradiation in locally advanced rectal cancer.

**sphincter sparing end points**

The better the response to chemoradiation, the more likely treatment will facilitate sphincter sparing options. So, whether the patient has a stoma or not, could be an end point of chemoradiation. Impressive results appear to have been achieved in phase II studies with chemoradiation [82, 83] and long-term follow-up has confirmed an excellent outcome if there is marked shrinkage of the distal margin of the tumour [84]. Preoperative chemoradiation appears to offer a 10% [85] to 20% better chance [80] for achieving sphincter sparing procedures. Whether a surgeon attempts sphincter sparing surgery depends on many factors including tumour size, location and accessibility, the level of training and experience of the surgeon, and the surgeon’s philosophy regarding risk and what to do with a tumour that prior to therapy appeared to require an AP excision and that after shrinkage by radiation and/or chemotherapy, now appears amenable to a sphincter sparing procedure: do you cut where the tumour was or do you cut where the tumour currently appears to be?

Randomised stratification by the surgeon may be appropriate, particularly if there is a large variation between participating surgeons in their level of training, expertise and philosophy on how to handle tumours that respond to preoperative therapy. Future randomised trials that expect to have a significant proportion of patients undergoing sphincter sparing procedures will need to incorporate standard measures of sphincter function.

**validation of rectal cancer surrogate end points**

There has been much controversy over the meaning of surrogate end points, and their statistical validation. Temple has defined a surrogate end point of a clinical trial as ‘a laboratory measurement or a physical sign used as a substitute for a clinically meaningful end point that measures directly how a patient feels, functions or survives’. Changes induced by a therapy on a surrogate end point are expected to reflect changes in a clinically meaningful end point [86]. This definition implies that it is not sufficient for an early end point (say, achievement of a negative resection margin) to be correlated with a later end point (say, being disease-free at 3 years) for the former to qualify as a surrogate for the latter.

For a surrogate end point to be valid, two conditions needed to be fulfilled [87, 88]: first, the surrogate and the true end points must be highly correlated ('individual-level surrogacy') and secondly, the effects of treatments on the surrogate and the true end points must be highly correlated ('trial-level surrogacy').

It could be worthwhile to adopt this approach to show that the circumferential resection margin provides an acceptable surrogate for 3-year DFS. The first condition implies that patients with R0 have a better 3-year DFS than those with R1. Based on the published literature from both randomised and non-randomised trials, this seems to be the case.

Figure 1 shows disease-free survival curves for a series of patients achieving R0 and R1/2 resection and reflects the impact of a positive CRM on disease-free survival. The study [33]
comprised 150 consecutive patients with unresectable rectal adenocarcinoma at Mount Vernon Hospital. All patients underwent preoperative CRT between 1994 and 2002 to a dose of 45 Gy in 25 fractions over 5 weeks using 5-FU and low-dose folinic acid followed by total mesorectal excision (TME).

Patients who achieve R0 have a much better prognosis, with a 3-year disease-free survival of about 52%, while patients who are R1 or R2, with a mere 9% 3-year disease-free survival. Hence the first condition of a strong association between the surrogate and the final end point seems to hold.

The second condition is far more challenging, as it can only be tested in the context of a series of randomised trials, which implies recourse to meta-analysis based on individual patient data. A few examples of such meta-analyses have recently been presented: in advanced colorectal cancer, the effects of treatment on tumour response were shown to correlate poorly with the eventual effects of these treatments on overall survival [84], while in early colorectal cancer, the effects of treatment on 3-year DFS were shown to correlate very well with the effects of treatment on 5-year survival [18]. In the case at hand, interest would focus on showing that the effects of a randomised therapy on R0 and on 3-year DFS were highly correlated.

The curves of Figure 1 suggest the magnitudes of benefit that might reasonably be expected in terms of 3-year disease-free survival if the proportion of patients with R0 could be increased, under the assumption that circumferential resection margin was a perfect surrogate for 3-year DFS. If an experimental chemoradiation regimen could increase the proportion of R0 from 63% as in the Mount Vernon series to 83% (a 20% benefit), Figure 1 suggests that the 3-year DFS might increase from 36% as in the Mount Vernon series to 45% (a 9% benefit). However, such predictions rest on the strong assumption that increases in the proportion of patients with R0 will be reflected in increases in 3-year DFS, and that such increases will follow the cross-sectional patterns shown on Figure 1. These predictions must be validated using data from randomised trials on the benefits of experimental therapies on the proportion of patients with R0 and on the proportion of patients disease-free at 3 years. If these benefits were highly correlated, R0 would be claimed to be a valid surrogate to predict the benefits of experimental therapies on 3-year disease-free survival [19].

The same approach could be used for any pair of the end points discussed above to investigate the surrogacy between early indicators of treatment effect and the late outcomes that are most clinically meaningful. The number of trials and patients required to carry out formal validations is a topic of current research, but it seems unlikely that any statistically reliable claims of surrogacy can be made with less than a few thousand patients. The meta-analyses mentioned above had included, respectively, 25 trials with some 3791 patients [87], 13 trials with 4352 patients [89] and the second 15 trials with some 13 000 patients [18]. Documenting the end points discussed above in sufficiently large numbers of randomised trials will allow analysts to perform a formal quantitative validation meta-analysis when data are mature in all trials.

conclusions

Alternative study end points should be measurable, sensitive, easy to interpret and clinically relevant, reflecting a tangible benefit to the patient. Disease-free survival currently remains the best (because it is relatively quick) primary end point in designing randomised phase III studies of preoperative chemoradiation in rectal cancer, although it is necessary to control for postoperative adjuvant chemotherapy. Disease-free survival is clearly an objective end point, which is important to the patient and clinically meaningful. There is also sufficient data from previous trials to calculate sample sizes required to detect differences of clinical interest.

The best secondary end point is the CRM, because it predicts the risk of both local recurrence and metastatic disease, but requires considerable quality control for surgery and pathology. We recommend the examination of the circumferential margin (CRM ≥1 mm) should be mandated in all future studies in rectal cancer. Hopefully, future randomised controlled trials will allow us to explore the effect of different chemoradiation schedules on this immediately observed end point, disease-free survival, and eventually overall survival.

references


