This is a repository copy of The D Prefix: Toward a Reproducible Validated Alternative End Point in Rectal Cancer.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/95180/

Version: Accepted Version

Article:
Glynne-Jones, R, West, NP orcid.org/0000-0002-0346-6709 and Quirke, P orcid.org/0000-0002-3597-5444 (2015) The D Prefix: Toward a Reproducible Validated Alternative End Point in Rectal Cancer. Diseases of the colon and rectum, 58 (6). pp. 613-616. ISSN 0012-3706

https://doi.org/10.1097/DCR.0000000000000356

© 2015, The American Society of Colon and Rectal Surgeons. This is an author produced version of a paper published in Diseases of the Colon and Rectum. Uploaded in accordance with the publisher's self-archiving policy. The final publication is available at Springer via http://dx.doi.org/10.1097/DCR.0000000000000356

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher’s website.

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.
The d-prefix: towards a reproducible validated alternative endpoint in rectal cancer.

R Glynne-Jones, N P West, P Quirke

Rob Glynne-Jones FRCP, Mount Vernon Centre for Cancer Treatment, Northwood, Middlesex, United Kingdom

Nicholas P. West PhD FRCPath and Philip Quirke PhD FRCPath, Pathology and Tumour Biology, Leeds Institute of Cancer and Pathology, School of Medicine, University of Leeds, Leeds, United Kingdom

Running title: The d-prefix in rectal cancer.

Corresponding Author: Dr R Glynne-Jones,
Tel: 44 (0) 1923 844767,
Fax: 44 (0) 1923 844 138,
E-mail: rob.glynnejones@nhs.net

Acknowledgements
NPW and PQ are supported by grants from Yorkshire Cancer Research, Harrogate, UK. NPW is also the recipient of a National Institute for Health Research Academic Clinical Lectureship, an Academy of Medical Sciences starter grant for clinical lecturers and a Pathological Society of Great Britain & Ireland Career Development Fellowship. PQ is also supported by the Experimental Cancer Medicine initiative.

Abstract
Recently the National Surgical Adjuvant Breast and Bowel Project investigators proposed the neoadjuvant rectal cancer score (NAR score) using ultimate pathological nodal stage (pN) and downstaging of T stage (ie cT – pT) as an early alternative endpoint. We recommend individual patients should have a ‘d’ prefix denoting the interval between the start of treatment and its assessment, local excision or definitive surgical procedure- denoted in days e.g ypd10 T2 ypN0 or ypd119 T2 ypN0, which would allow researchers to compare these results both within and across these groups in different chemoradiation/radiation studies.
Keywords: Locally advanced rectal cancer, total mesorectal excision, neoadjuvant chemotherapy, preoperative radiotherapy, chemoradiotherapy, downstaging, tumour regression.

For Debate

Fluoropyrimidine-based preoperative chemoradiotherapy (CRT) followed by total mesorectal excision is considered the current standard in the management of patients with clinically defined “locally advanced rectal cancer” (LARC). Current NCCN guidelines recommend chemoradiation for all patients clinically staged as cT3/T4. The recent publication of the early surgical endpoints in the National Surgical Adjuvant Breast and Bowel Project NSABP-R04 trial comparing preoperative radiation therapy and Capecitabine with or without Oxaliplatin with preoperative radiation therapy and continuous intravenous infusion of 5-Fluourouracil with or without Oxaliplatin,1 and presentations of preliminary data2 raise important points regarding the difficulty of finding an early endpoint to gauge the efficacy of CRT when evaluating different or novel regimens.,

The gold standard endpoint in clinical trials of chemotherapy and radiotherapy in rectal cancer remains overall survival. Although reliable and easy to measure, this endpoint takes years to observe. However, an alternative early endpoint, which would provide an accurate earlier assessment of treatment effects, would be useful. Neoadjuvant chemoradiotherapy achieves significant tumor downstaging/downsizing with pathological complete response (pCR) in up to 30% of cases in some series where early cancers have been included. Following completion of CRT, both individual series , population studies and a meta-analysis3 all show that longer intervals up to a maximum of 12 -15 weeks appear associated with an increased chance of achieving a pCR at surgical resection, and counter-intuitively outcomes may also improve in terms of a significant reduction in 3-year local recurrence rate (1.2% vs. 10.5%, p = 0.04).4 However, further extensions of this interval do not appear to benefit the patient.5,6 Patients with rectal cancer, who achieve a pCR or near pCR, fare consistently better than the patients who fail to do so.7 In contrast, patients with no evidence of response fare badly. Yet how to define best this response/lack of response and which method in terms of clinical measurement, imaging or pathology to measure size, volume, tumour cell-density, T-stage and N-stage downstaging, regression or residual functional activity remains a controversial issue.

Alternative early study endpoints are important because rapid methods to define and quantify the clinical utility of novel strategies such as dose escalation of radiotherapy or the integration of new drugs, and to be able to compare different strategies within clinical trials using CRT would be a major advantage. These endpoints should be objective, measurable, sensitive, easy to interpret and clinically relevant, reflecting a tangible benefit to the patient.

In contrast, the currently accepted late endpoints demand a long period from the end of recruitment to primary efficacy analysis for endpoints such as 5 year local recurrence or disease free survival (DFS) or overall survival (OS). Long term follow-up has the advantage of capturing long-term late toxic effects and second malignancies which may differ between treatments when a large proportion of patients are cured. Yet, this protracted period may also allow patients who fail later in the trial, to receive potentially more effective treatments and
survive longer. Outcome could therefore relate to both the intensity of preoperative radiotherapy or chemoradiation, the use of varying postoperative adjuvant chemotherapy or the eventual availability of more effective palliative treatments. Also interpatient heterogeneity with diversity in the phenotypic, epigenetic, and gene expression patterns between different rectal cancers, and intratumour heterogeneity within the same individual may blur the assessment of efficacy of any treatment.

The Union for International Cancer Control (UICC) uses a TNM classification to capture the extent of a cancer. There is a process for revising the TNM in the light of new knowledge. However, the pathological T stage (pT), and the pathological N stage (pN) reflect both the initial original preoperative stage but may also be affected/altered/modified by any preoperative therapy if effective - when it is given a ‘y’ prefix. As is pointed out - pCR – i.e. ypT0 ypN0 - only captures a small proportion of the patients, and is not discriminatory for prognosis in the group who fail to achieve pCR.

The rate of pCR within a CRT study reflects the initial proportion of early stage cancers as pCR after 5FU-based CRT is to some extent dependent on stage. In one pooled analysis the rate of pCR was 58% for T1, 28% for T2, 16% for T3 and 12% for T4 cancers respectively. PCR is also largely dependent on the degree of histopathological sampling, which is infrequently standardised within study protocols leading to problems with comparing different studies across the literature. If the entire area of scarring is blocked out and examined at multiple levels then the pCR rate will be significantly lower than if the cases are less intensively sampled.

An analysis of pooled data from five large European randomized clinical trials for locally advanced rectal cancer recommended the concept of nomograms to predict outcome. They suggested that 2-year DFS could be considered as an intermediate end point in future trials and this is being used in the current UK national CRT trial ‘Aristotle’. The nomograms use data from the preoperative assessments and postoperative histology and treatment, and weight different items with more or less importance. It is therefore likely that a more inclusive composite endpoint will prove useful in this setting. The potential advantage of a valid early composite endpoint is that it will not be influenced by the potential variety of postoperative treatments. The disadvantage is that practices for histopathological reporting vary on both sides of the Atlantic and even within Europe although we are rapidly acquiring a common language. If validated these alternative endpoints would hopefully enable the sample size of any study to be reduced, and hence also the duration of the trial. With this in mind, the NSABP investigators have suggested the neoadjuvant rectal cancer score (NAR score) using ultimate pathological nodal stage (pN) and downstaging of T stage ( ie cT – pT) , which they based on relative weights suggested for each item by the nomograms cited above.

\[ NAR = \frac{[5pN-3 \text{ (cT-pT)} + 12]^2}{9.61} \]

where cT = clinical T stage (1,2,3 or 4)
pT = pathological T stage (1,2,3 or 4)
pN = pathological nodal status (0,1 or 2)
captures the difference in T staging and only the pathological nodal status (because they accepted that clinical nodal status is not robust. The NAR score uses values from 0 to 100 as a pseudocontinuous variable, where higher scores indicate a poorer prognosis providing a low, intermediate, and high risk of death based on tertiles. It is recommended that analyses based on the score should be stratified by cT. In an analysis using data from the phase III neoadjuvant NSABP R-04 trial, the NAR score proved better at predicting OS than pCR. Although others have confirmed that NAR can outperform pCR in predicting OS, even NAR remains an imperfect endpoint and may not suffice to predict a satisfactory level of the therapeutic efficacy.

In addition, the biology of tumours is heterogeneous and after CRT different intervals may be required to achieve complete clinical response (CCR), and the timing to best response may be partially dependent on tumour size. For this reason composite downstaging endpoints such as the NAR proposed from results of an overall comparison of the clinical and pathological staging in the NSABR04 trial may be the best way forward. Ceteris paribus as above, such endpoints may also depend on both the precision and homogeneity of the timing of assessment, by what means you got there (chemotherapy or radiotherapy) and the quality of mesorectal excision.

In the large neoadjuvant randomised phase III CRT trials, the median interval between completion of preoperative CRT and surgery is between 3 and 10 weeks – see Table 1. Hence the interval may influence the selection of patients for adjuvant therapy on the basis of pathological features in the resected specimens, and alter the later outcomes. The NSABP R04 performed surgery also at 4 -6 weeks. So their results and the NAR scale are consistent with many of the randomised trials. However, up to 10 weeks is often reported among other published non-randomised clinical trials, and this interval is extending because of the influence of the Habr Gama ‘watch and wait’ data.

A retrospective analysis of the Swedish rectal cancer trial showed down-staging is observed after short course preoperative radiotherapy (SCPRT), when the interval to surgery is extended >10 days from the first fraction of radiotherapy. Further extension of the interval following SCPRT in surgery to at least six weeks, allows significant downstaging but probably not to the same extent as long course chemoradiation. To this end some have argued that the interval from the start of treatment to the time of surgery should be similar if SCPRT and CRT are to be compared.

In the Stockholm III trial when the first 400 patients were evaluated there was an increase in downstaging in those patients where the interval to surgery was extended to 8 weeks compared immediate surgery within 7-10 days. The pCR increased from 2% to 13% (p=0.001). A Dutch retrospective study supports this. Yet, downstaging ‘per se’ may not influence DFS or OS.

The Polish trial and the TROG trial demonstrate that the interval to pathological assessment influences the rate of response. The amount of downstaging and the pCR is very different in SCPRT with immediate surgery and CRT after a delayed interval, and the lack of nodal downstaging in the SCPRT arm has no impact on outcome - whereas it does in the CRT arm. Also the ACCENT database shows that the surrogacy of DFS for OS, is still present but at a much longer timeframe (OS at 7 years) for the addition of oxaliplatin to 5FU based chemotherapy. The outcomes for a pCR or persistent positive nodes (ypN+) after RT may be different to the outcomes after CRT.
The only published randomised controlled trial comparing two different interval lengths is the Lyon R90-01 trial of preoperative radiotherapy which randomised between a short interval or delayed surgery from the completion of radiotherapy (i.e. within 2 weeks, or 6-8 weeks). This study elegantly demonstrated that a longer interval increases the pCR rate.\textsuperscript{23,24} In the Dutch Surgical Colorectal audit evaluable patients who underwent preoperative CRT for rectal cancer between 2009 and 2011 were evaluated to determine the influence of the interval between radiotherapy and surgery, which was calculated from the start of radiotherapy.\textsuperscript{7} In this study an interval of 15-16 weeks after the start of CRT resulted in the highest pCR rate (18.0 per cent; \( P = 0.013 \)), with an independent association (hazard ratio 1.63, 95 per cent confidence interval 1.20 to 2.23). In a recent Polish study, 154 patients were randomized to SCPRT with surgery either 7–10 days or 4–5 weeks after the end of RT.\textsuperscript{25} More downstaging was seen after a longer interval. Hence, several alternative sequencing approaches have been examined.

However, for the NAR score to work in future trials the timing of surgery would have to be identical for every individual patient, and the two arms with and without oxaliplatin would have to be compared in terms of their eventual outcomes and be confirmed as similarly prognostic to validate NAR as a surrogate endpoint. For this reason, similarly to the prefix “y” used to denote pTN stage post preoperative chemotherapy or radiotherapy, we would propose the use of a novel prefix ‘d’ which is intended to define the timing of surgery in weeks or even refined to days. This ‘d’ defines the precise timing so that downstaging is captured histologically as a snapshot. One option if only chemoradiation is considered is to define the timing in weeks after the completion of surgery. However, if SCPRT and a delay is used, then it may be more appropriate to define the timing from the start or the first fraction of treatment. This could also apply as the ‘y’ prefix does to neoadjuvant chemotherapy alone.

Hence in summary, if we are to more effectively define and compare the benefits of different preoperative therapies the time has come to be more precise in the reporting of the interval between the start of treatment and surgery. We recommend that individual patients would have a ‘d’ prefix denoting the interval between the start of treatment and its assessment, local excision or definitive surgical procedure. This interval should be denoted in days, and might vary between 10 and 119 days e.g ypd10 T2 ypN0 or ypd119 T2 ypN0. Thus the overall trial could be analysed according to a median/mean ‘d’ score.

The NSABP results show that we need to provide a common metric across a range of clinical stage and tumour size, reducing the noisy data of a range of measures currently being used in clinical research, which would eventually allow researchers to compare these results both within and across these groups in different studies. The ‘d’ prefix is a first step.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Numbers</th>
<th>pCR%</th>
<th>Loc Recurrence</th>
<th>Interval from start of CRT</th>
<th>Interval from end of CRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACA0/ARO/AIO-94 (Sauer 2004)</td>
<td>394</td>
<td>8%</td>
<td>6%</td>
<td>9-11 weeks</td>
<td>4-6 weeks</td>
</tr>
<tr>
<td>POLISH (Bujko 2004, 2006)</td>
<td>157</td>
<td>13.4%</td>
<td>15.6%</td>
<td>9-11 weeks</td>
<td>4-6 weeks</td>
</tr>
<tr>
<td>Study</td>
<td>Designation</td>
<td>n</td>
<td>Pathological Complete</td>
<td>Time Between Chemoradiotherapy and Surgery</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>-----</td>
<td>-----------------------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>FFCD 9203</td>
<td>(Gerard 2006)</td>
<td>375</td>
<td>11%</td>
<td>8-15 weeks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8%</td>
<td>3-10 weeks</td>
<td></td>
</tr>
<tr>
<td>EORTC 22921</td>
<td>(Bosset 2005, 2006, 2013)</td>
<td>505</td>
<td>11.7%</td>
<td>8-15 weeks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8%</td>
<td>3-10 weeks</td>
<td></td>
</tr>
<tr>
<td>TROG 01-04</td>
<td>(Ngan 2012)</td>
<td>163</td>
<td>15%</td>
<td>9-11 weeks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.4%</td>
<td>4-6 weeks</td>
<td></td>
</tr>
<tr>
<td>NSABP R-03</td>
<td>(Roh 2009)</td>
<td>123</td>
<td>15%</td>
<td>Up to 13 weeks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.7%</td>
<td>no later than 8 weeks</td>
<td></td>
</tr>
</tbody>
</table>


