Pathologic complete response and disease-free survival are not surrogate endpoints for 5-year survival in rectal cancer: an analysis of 22 randomized trials

Fausto Petrelli, Karen Borgonovo, Mary Cabiddu, Mara Ghilardi, Veronica Lonati, Sandro Barni

Oncology Department, UO Oncologia, ASST Bergamo Ovest, 24047 Treviglio, BG, Italy

Contributions: (I) Conception and design: All authors; (II) Administrative support: All authors; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Fausto Petrelli, MD. Oncology Department, Oncology Division, ASST Bergamo Ovest, Piazzale Ospedale 1, 24047 Treviglio, BG, Italy. Email: faupe@libero.it.

Background: We performed a literature-based analysis of randomized clinical trials to assess the pathologic complete response (pCR) (ypT0N0 after neoadjuvant therapy) and 3-year disease-free survival (DFS) as potential surrogate endpoints for 5-year overall survival (OS) in rectal cancer treated with neoadjuvant (chemo)radiotherapy (CT)RT.

Methods: A systematic literature search of PubMed, EMBASE, the Web of Science, SCOPUS, CINAHL, and the Cochrane Library was performed. Treatment effects on 3-year DFS and 5-year OS were expressed as rates of patients alive (%), and those on pCR as differences in pCR rates ($\Delta pCR$). A weighted regression analysis was performed at individual- and trial-level to test the association between treatment effects on surrogate ($\Delta pCR$ and $\Delta 3yDFS$) and the main clinical outcome ($\Delta 5yOS$).

Results: Twenty-two trials involving 10,050 patients, were included in the analysis. The individual level surrogacy showed that the pCR% and 3-year DFS were poorly correlated with 5-year OS (R=0.52; 95% CI, 0.31–0.71; P=0.002; and R=0.60; 95% CI, 0.36–1; P=0.002). The trial-level surrogacy analysis confirmed that the two treatment effects on surrogates ($\Delta pCR$ and $\Delta 3yDFS$) are not strong surrogates for treatment effects on 5-year OS% (R=0.2; 95% CI, −0.29–0.78; P=0.5 and R=0.64; 95% CI, 0.29–1; P=0.06). These findings were confirmed in neoadjuvant CTRT studies but not in phase III trials were 3-year DFS could still represent a valid surrogate.

Conclusions: This analysis does not support the use of pCR and 3-year DFS% as appropriate surrogate endpoints for 5-year OS% in patients with rectal cancer treated with neoadjuvant therapy.

Keywords: Rectal cancer; pathologic complete response (pCR); disease-free survival (DFS); surrogate endpoints; overall survival (OS)

Submitted Jul 16, 2016. Accepted for publication Sep 01, 2016.
doi: 10.21037/jgo.2016.11.03
View this article at: http://dx.doi.org/10.21037/jgo.2016.11.03

Introduction

Neoadjuvant (chemo)radiotherapy (CT)RT is the gold standard of care for locally-advanced rectal cancer. The aim of treatment is to decrease local recurrence and improve: R0 surgery with a total mesorectal excision; and disease-free survival (DFS) and overall survival (OS). Either a short course of RT with immediate surgery or a prolonged course of (5-fluorouracil-based)-CTRT with delayed surgery are acceptable options for the preoperative treatment of rectal cancer, with or without adjuvant therapy (1,2). Overall, the effect of the addition of a short course of neoadjuvant RT prior to planned surgery is similar to that of a prolonged
course of CTRT in terms of survival, local and distant recurrences, and R0 resection, with more pathologic complete responses (pCRs) with combination therapy (3). A ypT0N0 stage pCR at the histologic examination after CTRT and surgery is commonly associated with better outcomes compared to non pCR patients, with less local and distant failure (4). The addition of multidrug regimens to standard RT has conferred increased toxicity, but has not led to a better rate of pCR in phase III trials. In particular, the addition of oxaliplatin to neoadjuvant 5FU-based CTRT only modestly improved the overall pCR rate (5).

The relationship between the response to neoadjuvant (CT)RT and the prognosis of patients with rectal cancer does not imply that pathological down-staging (e.g., pCR) is also a surrogate for treatment efficacy (OS). De facto, the demonstration that the pCR is a valid surrogate marker of the efficacy of systemic therapy on survival would encourage the use of primary systemic treatment to expedite the development of new systemic therapies with randomized neoadjuvant trials in locally-advanced rectal cancer (6). Furthermore, in colon cancer, DFS at 2 and 3 years is a good surrogate for OS at 5–6 years in trials of adjuvant CT (7-10).

To assess the roles of the pCR and DFS as potential surrogates of true clinical outcomes at the trial level, we performed a systematic literature search and a trial-based meta-analysis of randomized controlled trials comparing different neoadjuvant treatments that had available data on both the observed rates of pCR and 3-year DFS% and 5-year OS% outcomes, respectively. The aim of this study was to assess whether the treatment effects on the pCR and DFS are able to predict the treatment effects on OS.

Methods

We performed a literature-based analysis of randomized controlled trials of neoadjuvant RT or CTRT for rectal cancer. The primary objective of the meta-analysis was the individual and trial-level validation of the pCR% and 3-year DFS% as surrogate endpoints of the effect on 5-year OS% of neoadjuvant therapy in rectal cancer (e.g., evaluating whether the treatment effect on the pCR%, termed $\Delta^{pCR\%}$, and difference ($\Delta$) in 3-year DFS% allows the size of the effect on the main clinical endpoint, namely $\Delta$5-year OS%, to be predicted).

Literature search and study selection

A systematic literature search was conducted of PubMed, the Web of Science, SCOPUS, CINAHL, the Cochrane Library, and Embase up to August 2015. The search terms included “rectal cancer” or “rectal carcinoma”, “neoadjuvant” or “preoperative”, “chemotherapy” or “chemoradiotherapy” or “chemoradiation” or “radiotherapy”, and “randomized” or “randomised”.

The search was limited to phase II–III clinical trials published in the English language involving $\geq$100 patients. Two researchers (FP and AC) reviewed each abstract and text against the study inclusion and exclusion criteria. Studies were included if they: evaluated RT or CTRT (plus or minus adjuvant CT) as neoadjuvant therapy for rectal cancer followed by radical surgery; and reported both 5-year OS and either 3-year DFS (or progression/relapse free survival, or time to treatment failure provided they reported similar events of DFS) or the pCR% clearly defined as the % of ypT0N0 stages after preoperative therapy. Retrospective or prospective case series and phase I studies were excluded. Trials randomizing operated patients to adjuvant vs. no adjuvant therapy were considered provided they were all treated with neoadjuvant RT or CTRT, and included all patients who underwent preoperative treatment. In the event that a study was published in multiple articles or abstracts, the most recent data were used.

Data extraction

For each included study, data were extracted for study design, year of publication, sample size per treatment arm, and treatment schedule. Data on the pCR%, 5-year OS%, and 3-year DFS% were also collected. Rates of 3-year DFS and 5-year OS were captured from the reported Kaplan-Meier (KM) curves (11-13). Only in the case KM estimates were not presented, they were extracted from the article.

Statistical analysis

The statistical analysis consisted of a weighted linear correlation between the primary endpoint (5-year OS) and the candidate surrogates pCR%, and 3-year DFS. Analysis was weighted to the effective sample size at the time point considered (KM estimates of 3-year DFS and 5-year OS): number of events prior to the time point plus the number of patients at risk at the time point.

In particular, two correlations were calculated between the summary statistics to determine surrogacy according to methods previously reported (14-16). The first approach, termed individual-level surrogacy, computed the association...
between pCR% and 3-year DFS%, the potential surrogate endpoints, and 5-year OS%, for each included arm. The correlation was evaluated over all the treatment arms and is described as R (Pearson correlation coefficient). The R-squared ($R^2$) determination coefficient (the proportion of variability in OS explained by the variability of the surrogate endpoint) was also presented (17-19). The second approach, termed trial-level surrogacy, assessed the association between the reported treatment effects on a surrogate ($\Delta_{3yDFS}$ and $\Delta_{pCR}$), and those on OS ($\Delta_{5yOS}$), which is the main endpoint. A strong correlation ($R>0.8$) would be consistent with surrogacy for OS (20). As a sensitivity analysis, we explored the surrogacy of the pCR and DFS in CTRT containing arms and phase III studies only. Both analyses were weighted on the sample size of each trial included.

As the number of included trials was small, we applied the non-parametric bootstrap re-sampling method (using 10,000 bootstrap samples), weighted for the sample size of each trial, to construct the 95% confidence intervals (BCI) for all weighted correlation coefficients. All the reported P values correspond to 2-sided tests, and those that were less than 0.05 were considered to be statistically significant.

Analyses were performed with the NCSS 2007 software (version 07.1.21, released June 1, 2011).

**Results**

Following the systematic literature review, a total of 9,012 publications were analyzed (Figure 1), with 22 studies, published between 1999 and 2015, considered for inclusion in the final analyses (2,21-42). Most of the studies were randomized phase II (n=4) or III clinical trials (n=18).

The selected studies compared different CT backbones (n=6), different neoadjuvant treatments (RT vs. CTRT in n=5), and different strategies [neoadjuvant RT vs. surgery or neoadjuvant (CT)RT vs. adjuvant (CT)RT in n=5].

These studies involved 39 neoadjuvant treatment arms and 10,050 patients treated with some form of preoperative therapy (Table 1). There were between 50 and 924 patients with locally-advanced rectal cancer across the study treatment arms, and the reported 5-year OS rates ranged from 53% to 90% (median, 70.3%). In 4 arms, the 5-year OS% data were not available. The reported values for the 3-year DFS% ranged from 48% to 78% in n=22 arms (median, 70.5%). The pCR rates were presented in

---

**Figure 1** Flow diagram summarizing the strategy used to identify eligible studies.
<table>
<thead>
<tr>
<th>Author</th>
<th>Study/year</th>
<th>N (pts) (exp vs. ctr)</th>
<th>Neoadjuvant RT</th>
<th>Neoadjuvant CT (exp vs. ctr)</th>
<th>Adjuvant therapy</th>
<th>pCR% (exp vs. ctr)</th>
<th>Δ*pCR%</th>
<th>3Y-DFS% (exp-ctr)/Δ*3YDFS (%)</th>
<th>5Y-OS% (exp-ctr)/Δ*5YOS (%)</th>
<th>Median FU (months)/primary endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wong RTOG 0247</td>
<td>Phase II RCT/2012–2015</td>
<td>52 vs. 52</td>
<td>50.4 Gy/28 fx</td>
<td>CAPIRI vs. CAPOX</td>
<td>√</td>
<td>10.4 vs. 20.8</td>
<td>10.4</td>
<td>70-62/8</td>
<td>59-72–13</td>
<td>3.77–3.97 years/pCR</td>
</tr>
<tr>
<td>Rodel CAO/ARO/AIO-04</td>
<td>Phase III/2015</td>
<td>613 vs. 623</td>
<td>50.4 Gy/28 fx</td>
<td>5-FU + OXA vs. 5-FU</td>
<td>√</td>
<td>17 vs. 13</td>
<td>4</td>
<td>76-72/4*</td>
<td>78-80/–2</td>
<td>50/DFS</td>
</tr>
<tr>
<td>Sainato I-CNR/RT</td>
<td>Phase II/2014</td>
<td>334 vs. 321</td>
<td>45 Gy/25 fx</td>
<td>5FU + FA bolus gg 1–5, 29–33 (all pts)</td>
<td>Adj vs. no adj CT</td>
<td>18.6 vs. 17</td>
<td>1.6</td>
<td>70-70/0</td>
<td>66.9-67.9/–1</td>
<td>63.7/OS</td>
</tr>
<tr>
<td>Appelt</td>
<td>Phase II/2013</td>
<td>111 vs. 110</td>
<td>50 Gy/28 fx + brachy RT boost vs. 50 Gy/28 fx + EBRT boost</td>
<td>Oral UFT + FA</td>
<td>At discretion</td>
<td>18 vs. 18</td>
<td>0</td>
<td>62-67/–/–5</td>
<td>70.6–73.6/3</td>
<td>5.4 years/pCR</td>
</tr>
<tr>
<td>Saglam Istanbul R-01#</td>
<td>Phase II RCT/2014</td>
<td>76 vs. 77</td>
<td>45 Gy/25 fx</td>
<td>5FU 225 mg/m² d ic</td>
<td>Optional</td>
<td>1.3 vs. 1.3</td>
<td>0</td>
<td>74-74/0</td>
<td>76.5-74.2/2.3</td>
<td>56.8–59.3/local recurrences</td>
</tr>
<tr>
<td>Jeong**</td>
<td>Phase II/2014</td>
<td>170 vs. 170</td>
<td>45 Gy/25 fx + EBRT boost</td>
<td>CAP or 5FU + FA or UFT + FA or CAPIRI or CAPIRI + cetuximab</td>
<td>√</td>
<td>NR</td>
<td>NR</td>
<td>73-78/–5</td>
<td>83-88/–5</td>
<td>46–48/3Y DFS</td>
</tr>
<tr>
<td>Bosset EORTC 22921</td>
<td>Phase III/2006</td>
<td>506 vs. 505</td>
<td>45 Gy/25 fx</td>
<td>5FU + FA bolus gg 1–5, 29–33 (only pts in CTRT arms)</td>
<td>Randomization to adjuvant CT</td>
<td>13.7 vs. 5.3</td>
<td>8.4</td>
<td>NR/–</td>
<td>64.8-65.8/–1</td>
<td>5.4 years/OS</td>
</tr>
<tr>
<td>Gerard ACCORD 12</td>
<td>Phase III/2012</td>
<td>299 vs. 299</td>
<td>45 Gy/25 fx (arm 1) vs. 50 Gy/25 fx (arm 2)</td>
<td>CAP (arm 1) vs. CAPOX (arm 2)</td>
<td>At discretion</td>
<td>19.2 vs. 13.9</td>
<td>5.3</td>
<td>72-67.9/4.1</td>
<td>NR/–</td>
<td>36.8/pCR</td>
</tr>
<tr>
<td>Ngan Trans Tansman RTOG 01-04</td>
<td>Phase III/2012</td>
<td>163 vs. 163</td>
<td>50.4 Gy/28 fx (arm 1) vs. 25 Gy/5 fx (arm 2)</td>
<td>5FU 225 mg/m² d ic (arm 1)</td>
<td>√</td>
<td>NR</td>
<td>NR</td>
<td>68-74/–/–6</td>
<td>70-74/–4</td>
<td>5.9 years/local recurrences</td>
</tr>
<tr>
<td>Sauer CAO-ARO-AIO-94</td>
<td>Phase III/2004</td>
<td>415^</td>
<td>50.5 Gy/28 fx (arm 1)</td>
<td>5FU 1,000 mg/m² gg 1–5 week 1 &amp; 5 (arm 1)</td>
<td>√</td>
<td>8</td>
<td>NA</td>
<td>75/–</td>
<td>76/–</td>
<td>45.8/OS</td>
</tr>
<tr>
<td>Hofheinz</td>
<td>Phase III/2012</td>
<td>81 vs. 80^&lt;sup&gt;a&lt;/sup&gt;</td>
<td>50.4 Gy/5–6 weeks</td>
<td>CAP vs. 5FU 1,000 mg/m² gg 1–5 week 1 &amp; 5 (arm 1)</td>
<td>√</td>
<td>14 vs. 5</td>
<td>9</td>
<td>71-63/8</td>
<td>66-61/5</td>
<td>52/OS</td>
</tr>
<tr>
<td>Park</td>
<td>Phase III/2011</td>
<td>107^&lt;sup&gt;a&lt;/sup&gt;</td>
<td>50.4 Gy/25 fx</td>
<td>CAP</td>
<td>√ (+ randomization to adjuvant CTRT)</td>
<td>17</td>
<td>NA</td>
<td>77/NA</td>
<td>90/NA</td>
<td>52/3Y DFS</td>
</tr>
</tbody>
</table>

Table 1 (continued)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Author</th>
<th>Study/year</th>
<th>N (pts) (exp vs. ctr)</th>
<th>Neoadjuvant RT</th>
<th>Neoadjuvant CT (exp vs. ctr)</th>
<th>Adjuvant therapy</th>
<th>pCR% (exp vs. ctr)</th>
<th>∆&lt;sub&gt;3Y-DFS&lt;/sub&gt; (%)</th>
<th>3Y-DFS% (exp-ctr)/∆&lt;sub&gt;3Y-DFS&lt;/sub&gt; (%)</th>
<th>5Y-OS% (exp-ctr)/∆&lt;sub&gt;5Y-OS&lt;/sub&gt; (%)</th>
<th>Median FU (primary months)/endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roh NSABP R-03</td>
<td>Phase III/2009</td>
<td>123&lt;sup&gt;**&lt;/sup&gt;</td>
<td>45 Gy/25 fx (+ randomization to adjuvant CTRT)</td>
<td>5FU + FA bolus x6 weeks → 5FU + FA bolus weeks 1 &amp; 5</td>
<td>√</td>
<td>15</td>
<td>NA</td>
<td>70/NA</td>
<td>67/NA</td>
<td>NA/DFS &amp; OS</td>
</tr>
<tr>
<td>Braendengen</td>
<td>Phase III/2008</td>
<td>98 vs. 109</td>
<td>50 Gy</td>
<td>± 5FU bolus gg 1,2,11,12,21,22</td>
<td>5FU + FA in CTRT arm (permitted in RT arm)</td>
<td>16 vs. 7</td>
<td>9</td>
<td>65-48&lt;sup&gt;§&lt;/sup&gt;/13</td>
<td>66-53/13</td>
<td>61/5Y OS</td>
</tr>
<tr>
<td>Bujko</td>
<td>Phase III/2006</td>
<td>157 vs. 155</td>
<td>50 Gy/28 fx (arm 1) vs. 25 Gy/5 fx (arm 2)</td>
<td>5FU + FA bolus weeks 1 &amp; 5 (arm 1)</td>
<td>Optional</td>
<td>16.1 vs. 0.7</td>
<td>15.4</td>
<td>60-63/-3</td>
<td>NA/-</td>
<td>48/sphincter preservation</td>
</tr>
<tr>
<td>Mohiuddin RTOG-0012</td>
<td>Phase II RCT/2013</td>
<td>50 vs. 53</td>
<td>45.6 Gy/1.2 Gy fx bid + boost (arm 1) vs. 45 Gy/1.8 Gy fx bid + boost (arm 2)</td>
<td>5FU 225 mg/m&lt;sup&gt;2&lt;/sup&gt; d ic (arm 1) + 5FU 225 mg/m&lt;sup&gt;2&lt;/sup&gt; d ic + weekly OPT11 (arm 2)</td>
<td>-</td>
<td>30 vs. 26</td>
<td>4</td>
<td>NA/NA</td>
<td>61-75/-14</td>
<td>6.4–7 years/pCR &amp; toxicity</td>
</tr>
<tr>
<td>Pach</td>
<td>Phase II RCT/2012</td>
<td>77 vs. 77</td>
<td>25 Gy/5 fx (random to immediate vs. delayed surgery)</td>
<td>-</td>
<td>-</td>
<td>10.4 vs. 0</td>
<td>10.4</td>
<td>NA/-</td>
<td>73-63/10</td>
<td>86/recurrences and OS</td>
</tr>
<tr>
<td>Sebag Montefiore MRC CR07</td>
<td>Phase III/2009</td>
<td>674&lt;sup&gt;^&lt;/sup&gt;</td>
<td>25 Gy/5 fx (+ randomization to immediate CT)</td>
<td>-</td>
<td>At discretion</td>
<td>NR</td>
<td>-</td>
<td>70.3/-</td>
<td>77.5/-</td>
<td>48/local recurrence</td>
</tr>
<tr>
<td>Gerard FFCD 9203</td>
<td>Phase III/2006</td>
<td>367 vs. 375</td>
<td>45 Gy/25 fx</td>
<td>5FU + FA bolus weeks 1 &amp; 5 vs. no CT</td>
<td>√</td>
<td>11.4 vs. 3.6</td>
<td>7.8</td>
<td>NA/-</td>
<td>67.9-67.4/0.5</td>
<td>81/OS</td>
</tr>
<tr>
<td>Glehen Lyons R90-01</td>
<td>Phase III/2003</td>
<td>99 vs. 101</td>
<td>39 Gy/13 fx (random to immediate vs. delayed surgery)</td>
<td>-</td>
<td>-</td>
<td>7 vs. 14</td>
<td>7</td>
<td>NA/-</td>
<td>69-66/3</td>
<td>6.3 years/sphincter preservation &amp; local control</td>
</tr>
<tr>
<td>Peeters/Kapiteijn TME trial</td>
<td>Phase III/2001-2007</td>
<td>924&lt;sup&gt;**&lt;/sup&gt;</td>
<td>RT vs. TME surgery alone</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>NA/-</td>
<td>64.2/-</td>
<td>6 years/local control</td>
</tr>
<tr>
<td>Allegra NSABP R-04</td>
<td>Phase III/2015</td>
<td>1608</td>
<td>RT 45 Gy/25 fx</td>
<td>5FU ic or CAP ± OXA</td>
<td>Not known</td>
<td>19.5 vs. 17.8</td>
<td>1.7</td>
<td>NA/-</td>
<td>81.3-79/2.3</td>
<td>NA/locoregional control at 3 y</td>
</tr>
<tr>
<td>Wong RTOG 0247</td>
<td>Phase II RCT/2012-2015</td>
<td>52 vs. 52</td>
<td>50.4 Gy/28 fx</td>
<td>CAPEIRI vs. CAPOX</td>
<td>√</td>
<td>10.4 vs. 20.8</td>
<td>10.4</td>
<td>70-62/8</td>
<td>66-46/20</td>
<td>3.77-3.97 years/pCR</td>
</tr>
</tbody>
</table>

*, statistically significant; **, randomized to open vs. laparoscopic surgery; †, the study investigated different timing of surgery (4 vs. 8 weeks after neoadjuvant therapy); ‡, randomization to neoadjuvant vs. adjuvant therapy but only neoadjuvant arm considered for the purpose of the study; ‡‡, only neoadjuvant arm; §, time to treatment failure; ∆<sub>3Y-DFS</sub>, relapse-free survival; ∆<sub>5Y-OS</sub>, progression-free survival. pCR, pathologic complete response; N, number; CT, chemotherapy; RT, radiotherapy; exp, experimental; ctr, control; 3Y-DFS, difference in 3-year disease-free survival rate; 5Y-OS, difference in 5-year overall survival rate; ref, reference; √, offered to all patients; pts, patients; NA, not applicable for neoadjuvant comparisons; UFT, uracil + tegafur; FA, folinic acid; ic, continuous infusion; 5FU, 5-fluorouracil; CAPOX, capecitabine + oxaliplatin; CAPEIRI, capecitabine + irinotecan; OXA, oxaliplatin; d, daily; NR, not reported; 3Y DFS, 3-year disease free survival; 5Y OS, 5-year overall survival; fx, fractions.
Among a total of 39 treatment arms available, the values for the pCR%/5-year OS correlation were reported in n=30 of them. In the analysis of all the treatment regimens, the pCR% correlated weakly with OS (R=0.52; BCI 95% CI, 0.31–0.91; P=0.002; Figure 2). The R² values were 0.28 (P=0.002). The correlation between 3-year DFS/OS was available for n=22 arms and was similarly poor (R=0.6; BCI 95% CI, 0.36–1; P=0.002; Figure 3). R² was 0.37 (P=0.002).

Restricting the analysis to the phase III trials only (n=22 arms), the correlation of the pCR% with 5-year OS was moderate (R=0.60; P=0.002); the correlation of the 3-year DFS with 5-year OS was similar (R=0.61; P=0.01). In the studies that adopted CTRT treatment in all comparisons (n=17 arms and 19 arms for the DFS and pCR% analysis), the correlation of 3-year DFS/5-year OS was similar (R=0.66; P=0.0037). The correlation of the pCR% with 5-year OS was negligible (R=0.05; P=0.81).

Trial-level surrogacy

A total of 9 pairs of Δ₃yDFS and Δ₅yOS between the treatment arms were available in the randomized trials. The correlation between Δ₃yDFS and Δ₅yOS was 0.64 (BCI 95% CI, 0.29–1), and P=0.06. The correlations ΔpCR%/Δ₅yOS were available for 13 pairs of comparisons and R was 0.2 (BCI 95% CI, 0.29–0.78), and P=0.5 (Figure 4). The R² values were 0.41 and 0.04.

The slope of the regression equation, namely the estimated change in the Δ₅yOS per unit change in the rate of ΔpCR%, was 0.22, with a standard error of 0.33 [Δ₅yOS = (−1.08) + 0.22*ΔpCR%]. This means that a treatment associated with a 10% increase in ΔpCR% translated into an approximately (not

**Outcome surrogacy**

Among a total of 39 treatment arms available, the values for the pCR%/5-year OS correlation were reported in n=30 of them. In the analysis of all the treatment regimens, the pCR% correlated weakly with OS (R=0.52; BCI 95% CI, 0.31–0.91; P=0.002; Figure 2). The R² values were 0.28 (P=0.002). The correlation between 3-year DFS/OS was available for n=22 arms and was similarly poor (R=0.6; BCI 95% CI, 0.36–1; P=0.002; Figure 3). R² was 0.37 (P=0.002).

Restricting the analysis to the phase III trials only (n=22 arms), the correlation of the pCR% with 5-year OS was moderate (R=0.60; P=0.002); the correlation of the 3-year DFS with 5-year OS was similar (R=0.61; P=0.01). In the studies that adopted CTRT treatment in all comparisons (n=17 arms and 19 arms for the DFS and pCR% analysis), the correlation of 3-year DFS/5-year OS was similar (R=0.66; P=0.0037). The correlation of the pCR% with 5-year OS was negligible (R=0.05; P=0.81).

**Trial-level surrogacy**

A total of 9 pairs of Δ₃yDFS and Δ₅yOS between the treatment arms were available in the randomized trials. The correlation between Δ₃yDFS and Δ₅yOS was 0.64 (BCI 95% CI, 0.29–1), and P=0.06. The correlations ΔpCR%/Δ₅yOS were available for 13 pairs of comparisons and R was 0.2 (BCI 95% CI, 0.29–0.78), and P=0.5 (Figure 4). The R² values were 0.41 and 0.04.

The slope of the regression equation, namely the estimated change in the Δ₅yOS per unit change in the rate of ΔpCR%, was 0.22, with a standard error of 0.33 [Δ₅yOS = (−1.08) + 0.22*ΔpCR%]. This means that a treatment associated with a 10% increase in ΔpCR% translated into an approximately
significant) 2% increase in 5-year OS probability. Similarly, the slope of the regression equation, and the estimated change in the \( \Delta^{5\text{yOS}} \) per unit change in the \( \Delta^{3\text{yDFS}} \), was 0.51, with a standard error of 0.23 \( \Delta^{5\text{yOS}} = (-2.16) + (0.51)\Delta^{3\text{yDFS}} \). This means that a treatment associated with a 10% increase in 3-year DFS % translated into a non-significant 5% increase in the risk of 3-year chance of being progression-free or alive.

For the phase III and CTRT-only trials, the correlations of \( \Delta^{\text{pCR}} \) and \( \Delta^{3\text{yDFS}} \) with \( \Delta^{5\text{yOS}} \) were poor (R=0.78, P=0.11, and R=0.8, P=0.02 for the phase III trials; and R=-0.21, P=0.68, and R=0.17, P=0.71 for the CTRT trials, respectively).

**Discussion**

Rectal cancer patients with a pCR defined as no residual cancer found upon the histological examination of surgical specimens (ypT0N0) after CTRT have better long-term outcomes, less risk of developing local or distal recurrences, and improved survival. In particular, after neoadjuvant CTRT and delayed surgery, a pCR is obtained in 15–27% of patients (43). Patients obtaining a pCR have a 50% reduced risk of death and relapse, but they still portend a residual risk of local (2.8%) and distant (9%) metastases. In rectal cancer, which is a disease with a different biology and treatment approach compared to colon carcinoma, a formal validation of the surrogacy of the pCR and DFS is still lacking, and a demonstration of a correlation with OS would be required.

In the present analysis, with data extracted from a total of 22 trials, we estimated the correlation equation of the effect on the pCR and 3-year DFS% on the effect on the main outcome (5-year OS%). We observed that both the pCR and 3-year DFS are not candidates for surrogates of OS in rectal cancer studies. In particular, the \( R^2 \) results (0.02 and 0.42 for the 2 trial-level correlation analysis) suggest that the neoadjuvant effect on the pCR and 3-year DFS% are able to explain no more than 2% and 48%, respectively, of the effects detected on 5-year OS% in patients with rectal cancer.

Recently, Valentini et al. identified 2-year DFS more than pCR to be a good predictor of survival in a pooled analysis of five randomized European trials (44). They did not provide a formal surrogacy analysis, but did identify three risk classes of patient for whom reduced intensity treatment (in excellent and good prognosis subgroups) may be hypothesizable, as well as those with a poor prognosis (20% of total population) for whom more intensive/effective therapies do not lead to a definitive cure, with more efficacious therapies urgently awaited.

The question of the surrogacy of the pCR has arisen for other solid tumors with similar negative results (45,46). In our series, more intensive neoadjuvant schedules were offered in only three trials, and so a formal subgroup analysis was not performed. However, the results were similar in both larger phase III studies and those with concurrent CTRT in both comparison arms.

There could be several reasons for our findings, and this represents the main limitations of this analysis. First, this is a literature-based analysis, and more appropriate validation with individual patient data is necessary. Second, the relatively short follow-up for most trials did not potentially capture late recurrences, as shown in Valentini et al.’s analysis (5% more distant metastases were found at 10, compared to 5, years in patients who obtained a pCR). Third, some older trials with RT and surgery-alone arms, and with intrinsic technical issues related to radiation and surgical pathology accuracy, could have led to surprising results. Fourth, the randomized or non-choice of adjuvant CT in many trials could have diluted the final result. However, this is the first analysis that systematically evaluated the surrogacy of pCR and DFS with 5-year OS in rectal cancer through a systematic evaluation of 22 randomized trials of neoadjuvant therapy involving more than 10,000 patients. The analysis confirmed the negative findings of surrogacy for both intermediate endpoints in CTRT studies, but significant results for surrogacy were found in 5 large phase III trials for 3-year DFS endpoint.

With the possible influence of adjuvant and salvage therapies at relapse, the results of this trial-based meta-analysis indicated only a poor correlation between neoadjuvant treatment effects on the pCR and a moderate correlation of 3-year DFS% on 5-year OS%. The findings do not therefore support the use of these intermediate endpoints as surrogate endpoints of treatment efficacy in patients with locally-advanced rectal cancer treated with neoadjuvant-based therapy. New clinico-pathological and molecular biomarkers are potentially useful for predicting final outcomes. Among them, the NAR score has been developed based on cT, pT and pN pathological results (47,48). The score has been validated in the NSABP-R04 trial, and is emerging as a useful short-term surrogate clinical trial endpoint in rectal cancer study designs. This approach is undergoing trial level validation, and has already been adopted as a secondary and, possibly, primary endpoint in several ongoing phase I and II studies testing novel preoperative interventions in rectal cancer.
Further studies are needed to assess the surrogacy of the pCR in a small subgroup of patients with an excellent prognosis and for whom conservative surgery or the wait-and-see strategy can be options. In the meantime, due to the occurrence of late relapses and deaths identified in long-term follow-up observations in major phase III trials, 5-year OS should still remain the surrogate of a definitive cure for most patients.

Acknowledgements
None.

Footnote
Conflicts of Interest: The authors have no conflicts of interest to declare.

References
8. Sargent D, Shi Q, Yothers G, et al. Two or three year disease-free survival (DFS) as a primary end-point in stage III adjuvant colon cancer trials with fluoropyrimidines with or without oxaliplatin or irinotecan: data from 12,676 patients from MOSAIC, X-ACT, PETACC-3, C-06, C-07 and C89803. Eur J Cancer 2011;47:990-6.


