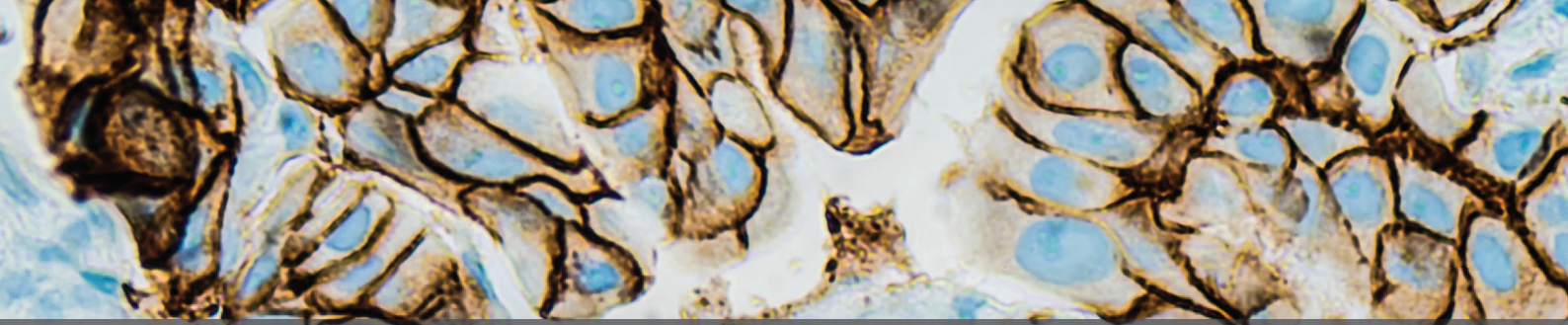

Neoadjuvant treatment of HER2 -positive breast cancer – pathological complete response (pCR) as a surrogate of long term outcomes in the context of regulatory guidelines and reimbursement recommendations

DOI:10.7365/JHPOR.2018.1.1

Authors:

Anna Panasiuk[1],
Dariusz Pawlik[1],
Małgorzata Budasz-Świderska[2],
Marcin Kaczor[3]

1 - Aestimo s.c., Kraków, Poland
2 - Roche Polska Sp. z o.o., Warszawa, Poland
3 - Uniwersytet Jagielloński Collegium Medicum, Kraków, Poland

Keywords:

review, surrogate endpoints, breast cancer, neoadjuvant treatment, pathologic complete response

How to cite this article?
Panasiuk A, Pawlik D, Budasz-Świderska M, Kaczor M, Neoadjuvant treatment of HER2-positive breast cancer – pathological complete response (pCR) as a surrogate of long term outcomes in the context of regulatory guidelines and reimbursement recommendations. J Health Policy Outcomes Res [Internet]. 2018 [cited YYYY Mon DD]; Available from: http://jhpor.com/Home/Article/2198 DOI:10.7365/JHPOR.2018.1.1

Abstract

Background: The vast heterogeneity of breast cancer (BC) patients, together with relatively long survival in the early stages of the disease, leads to methodological issues in designing clinical trials, particularly in case of therapies targeted for narrow patients subsets. That situation led to the need and search for early, reliable predictors of long-term outcomes.However, in countries adopting HTA methodology to reimbursement decisions non-final ("surrogate") endpoints may not be accepted as a measure of clinical benefit.

Methods: We searched PubMed and EMBASE for meta-analyses investigating association of pathologic complete response to BC neoadjuvant therapy with longer-term clinical outcomes. To be eligible analyses had to pool primary studies identified in a systematic manner. The results of the included studies were discussed in the context of current regulatory and reimbursement recommendations.

Results: The 4 included meta-analyses varied in methodological approaches, eligibility criteria and number of included patients. Three of them assessed predictive value of pCR in overall non-metastatic BC population and two – in HER2-positive subgroup. The existing differences in conclusions are consistent with the variability in methodological assumptions.

Conclusion: Significant heterogeneity of BC population and low number of studies showing substantial treatment effect on pCR could bias the results of "trial-level" analyses. In contrast, studies consistently show significant patient-level association of pCR with clinical outcomes, particularly strong for aggressive tumour subtypes. The existing EMA and FDA guidelines show how to make use of the existing evidence, in spite of its limitations, in a pursuit of satisfying the unmet medical needs.

Introduction

Long-term overall survival, together with signs and symptoms, safety and quality of life assessment constitute the most desirable set of endpoints for oncology clinical trials. The need for providing strong, direct evidence for "final" outcomes in the appraisal of new therapies for life-threatening diseases, preferably overall survival (OS), are emphasized in the European health technology assessment (HTA) guidelines.[1] In countries adopting HTA methodology to reimbursement decisions, providing definite evidence for OS improvement, that may be quite easily translated into quality-adjusted life-year gain estimate, seems to become a prerequisite for public funding for new oncology drug treatment, while other endpoints ("surrogate outcomes") may be rejected as not sufficient or not relevant as a measure of clinical benefit.[2, 3] Yet there are therapeutic areas where unmet clinical need coexists with significant obstacles in obtaining meaningful final endpoint assessment with reasonable time and costs. Development of new treatments for early-stage breast cancer (eBC) is a well-recognised area of such difficulties. Breast cancer (BC) is the most common female cancer in the world, in both more and less developed regions.[4] Advances in early diagnosis, related to an increasing implementation of population screening programmes, in addition to improvements in adjuvant systemic therapy led to substantial decrease in BC mortality in United States and western European Countries.[5, 6] At the same time, BC is a heterogeneous condition, where treatment outcomes may be influenced not only by clinical stage, but also by histological, molecular and functional features.[7] Correlations between differences in gene expression and response to treatment shown in recent studies add to the complexity of the disease and justify more personalised approach to BC therapy.[8]

The vast heterogeneity of BC patients, together with relatively long disease-free and overall survival in the early stages of the disease, inevitably leads to methodological issues in designing clinical trials for regulatory purposes, particularly in case of therapies targeted for narrow patients subsets. That situation led to the need and search for early, reliable predictors of long-term outcomes in eBC trials. Currently, an increase in pathologic complete response (pCR) rate after systemic, preoperative (neoadjuvant) treatment is an endpoint accepted by both the European Medicines Agency (EMA) and the U.S. Food and Drug Administration (FDA), though under additional, specified conditions and, in case of the FDA, sufficient only to support accelerated approval. The need for faster access to novel, improved therapies for high-risk/poor

prognosis BC patient subgroups was a key point justifying the use of pCR in the absence of definitive survival or event-free survival data, in spite of unavoidable uncertainty around surrogate-based efficacy assessment.[9, 10]

Despite positive, pragmatic attitude of the regulatory bodies towards early, pCR-based assessment of new therapies for high-risk BC, doubts may still be raised at the stage of post-approval appraisal on the national or regional level.[11] A quick review of recommendations issued by HTA agencies for neoadjuvant use of dual HER2-blockade (pertuzumab added to trastuzumab and chemotherapy) in HER2-positive, locally advanced or large operable BC between 2015 and 2017 shows considerable reluctance to acknowledge the clinical benefit of the treatment in the absence of direct evidence of the survival gain. In July 2015 the Canadian Agency for Drugs and Technologies in Health (CADTH) considered a significant increase in the pCR rate as not sufficient to conclude that the assessed technology resulted in a net clinical benefit compared with a current standard (trastuzumab and chemotherapy without pertuzumab addition).[12] The CADTH conclusion was justified by the statement that pCR "has not been validated as a surrogate for either event-free or overall survival"[12] (p1). In the same year German Institute for Quality and Efficiency in Health Care (IQWIG) issued a report for double HER2-blockade therapy in which pCR data were deemed "not patient-relevant" and as such excluded from assessment.[13] In January 2016, the IQWIG supplemented its assessment with additional data, regarding pCR results and their transferability to clinical outcome in a single clinical trial, sent by the Marketing AuthorisationHolder (MAH). Still, the IQWIG denied to accept those data, holding on to the assumption that only the occurrence of disease recurrence itself was the relevant endpoint, which could had not been substituted by its predictor.[14] In contrast to the CADTH and the IQWIG resolutions, in December 2016 the National Institute for Health and Care Excellence (NICE), on the base of the same clinical data, recommended use of neoadjuvant pertuzumab.[15] Despite "no reliable trial evidence of event-free or overall survival benefit"[15] (p18) and uncertainty about the exact relationship of pCR with long-term outcomes, the NICE committee came to the conclusion that "the complete disappearance of cancer in the breast and nodes was more likely to be associated with improved long-term outcomes than completely unrelated"[15] (p18). Hence, pCR was "more likely than not to have an association with longer-term survival"[15] (p18).

Puzzling as it may be, in the considerations of pCR value as an endpoint both CADTH and NICE referred to the same meta-analysis of the Collaborative Trials in Neoadjuvant Breast Cancer (CTneoBC) group[16] established by the FDA, that aimed at the assessment of pCR surrogacy for long-term clinical outcome. Moreover, the same CTneoBC work seems to be the key basis for pCR value appraisal that underlay EMA and FDA regulatory guidelines on this topic.[9, 10] Beyond the CTneoBC analysis, both regulatory documents refer also to other studies adding to the knowledge on pCR predictive value[9, 10], though none of them were discussed in CADTH, IQWIG or NICE recommendations.[12 – 15]

We aimed to address the existing inconsistency in approaches to selection and interpretation of data on pCR surrogacy in neoadjuvant BC treatments assessment by attempting an explicit, systematic approach to publications selection and review of their methodological assumptions. We searched for the best available evidence for pCR value as an surrogate outcome in HER2-positive, non-metastatic BC. To provide the relevant context for data interpretation, our review was not restricted to studies that focused solely on HER2-positive subgroup, but analyses conducted in broader eBC population, irrespective of molecular subtype, were also included.

Materials and methods

We searched PubMed and EMBASE in October 2017 for meta-analyses investigating association of pathologic complete response to neoadjuvant therapy in breast cancer with long-term clinical outcomes. Searches of databases were conducted using combinations of the following keywords: ((pathologic OR pathological) AND complete response[ti]) AND (breast cancer OR breast carcinoma) AND (pooled OR meta-analysis OR systematic). Reference lists were also searched for related articles. Eligible were analyses that pooled primary studies identified on the basis of systematic search and explicit selection criteria, that assessed association of pCR with overall, progression-free, event-free, disease-free or relapse-free survival in non-metastatic breast cancer population, unrestricted to any molecular subtype or restricted to HER2-positive patients, published in English or Polish. We excluded meta-analyses not based on systematic review or restricted to subpopulations other than HER2-positive. Purely narrative reviews and research not published as a full-text article was not eligible. From included papers we extracted selection criteria, required pCR definition (if stated), number of included studies, pooled population size, key methodological assumptions and results.

Results

A total of 65 potentially revelant articles were identified after the initial database search, of which 15 were analysed in full text. Four systematic-review based pooled analyses were included in the review, of Kong et al.,[17]

Table 1. Key eligibility criteria for primary data sources and other methodological features of included meta-analyses.

Feature	Kong 2011[17]	Berruti 2014[18]	CTneoBC[16]	Broglio 2016[19]
Searched period	Jan 1985 – Nov 2010	up to March 2013	Jan 1990 – Aug 2011	up to Dec 2014
Patients	BC	BC	BC, primary	stage I-III HER2-positive BC
Analysis in non-metastatic BC population, not selected by subtype	yes	yes	yes	no
Analysis in HER2-positive BC patients population	no	no	yes	yes
Treatment	neoadjuvant chemotherapy delivered before loco regional therapy of BC, no additional post-surgical adjuvant treatment (i.e. radiotherapy)	neoadjuvant chemotherapy or neoadjuvant HER2 targeted therapy and cytotoxic therapy	preoperative chemotherapy followed by surgery	neoadjuvant systemic therapy
Study design	-**	RCTs	clinical trials including at least 200 patients, with a median follow-up of at least 3 years	randomised or cohort studies (prospective or retrospective), studies that pooled trial participants and cohorts
Definition of pCR*	explicit	explicit and based on excision histology; choice of the pCR definition for the primary analysis followed hierarchical order: ypT0 ypN0, ypT0/is ypN0, ypT0, ypT0/is	ypT0 ypN0, ypT0/is ypN0, ypT0/is	publications were included regardless of definition of pCR, ypT0/is was used if available
Clinical outcomes	OS, DFS and RFS	OS and DFS (deemed equivalent to PFS, RFS and EFS)	OS and EFS	OS and EFS (used as an umbrella term for EFS, RFS, RcFS and DFS)
Other criteria for inclusion	detailed statistics reported for complete and partial responders; detailed outcomes for OS, RFS or RFS reported (not limited to p-values); published in English; primary results only	both pCR rates and survival outcomes reported; any post-surgical adjuvant treatments allowed; no language restrictions	data for pCR, EFS, and OS available	publications were included regardless of neoadjuvant regimen and definition of EFS
Patient-level analysis‡	yes, literature based	no	yes, IPD-based	yes, literature based
Trial-level analysis‡‡	no	yes, literature based	yes, IPD-based (RCTs only)	yes, literature based (RCTs only)
Number of included studies	16	29	12	38†
Number of included patients††	3776	14 641	11 955 (1989 HER2-positive)	5768 (all HER2-positive)

** ypT0 ypN0 - absence of invasive cancer and in-situ cancer in the breast and axillary nodes, ypT0/is ypN0 - absence of invasive cancer in the breast and axillary nodes, irrespective of ductal carcinoma in situ, ypT0 - absence of invasive cancer and in-situ cancer in the breast irrespective of nodal involvement; ypT0/is - absence of invasive cancer in the breast irrespective of ductal carcinoma in situ or nodal involvement; ** eligibility criteria for study design not given; ‡ chance/risk of clinical endpoint in responder vs in non-responder subgroup; ‡‡ percentage variability of treatment effect on clinical outcome explained by variability of treatment effect on pCR; † CTneoBC meta-analysis for HER2-positive subgroup was used as a primary source and counted as a single study; †† the table shows a total number of patients in pooled population of included study arms but the actual number of included patients differed for each variant of a meta-analysis*

Abbreviations: BC - breast cancer; DFS – disease-free survival; EFS – event-free survival; HER2 – human epidermal growth factor receptor 2; IPD - individual patient data; pCR – pathologic complete response; OS – overall survival; RCTs – randomised controlled trials; RcFS – recurrence-free survival; RFS - relapse-free survival

Berruti et al.,[18] the Collaborative Trials in Neoadjuvant Breast Cancer (CTneoBC) group[16] and Broglio et al.[19] Kong[17] and Berruti[18] assessed predictive value of pCR in general population of non-metastatic BC patients, regardless of molecular subtype. The analysis of FDA-established CTneoBC group was conducted both in general population and in prognostic subgroups, including HER2-positive group. The most up-to-date work of Broglio et al.[19] was restricted solely to HER2-positive BC patients. Table 1. summarizes key methodological features of each of the analyses included.

The earliest meta-analysis[17] included 3776 BC patients, treated with neoadjuvant chemotherapy in 16 clinical trials. Included patients received various chemotherapies, but the vast majority of patients received a regimen that consisted of an anthracycline and/or taxane with other agents. To avoid confounding related to post-surgical treatment, studies in which additional adjuvant treatment was delivered (i.e. radiotherapy) were excluded. Pooled odds ratio (OR) with 95% confidence interval (CI) of achieving beneficial endpoint (overall, relapse-free or disease-free survival) in patients who achieved pCR (responders) versus those who did not achieve pCR (non-responders) after neoadjuvant treatment served as an indicator of pCR predictive value. Meta-analyses showed positive, significant association of pCR with OS [OR = 3.44 (95% CI: 2.45 to 4.84); $p < 0.00001$], as well as with disease-free survival (DFS) [OR = 3.41 (95% CI: 2.54 to 4.58); $p < 0.00001$] and relapse-free survival (RFS) [OR = 2.45 (95% CI: 1.59 to 3.80); $p < 0.0001$]. There was no separate analysis for HER2-positive subgroup. The results led to conclusion that pCR was a prognostic indicator for relapse-free, disease-free and overall survival and suggested that patients achieving pCR after neoadjuvant chemotherapy had favourable outcomes.

The next included study[18] was based on meta-regression of 29 randomised controlled trials (RCTs), jointly comprising of 59 arms. A total of 14 641 patients were treated with neoadjuvant chemotherapy or neoadjuvant anti-HER2 targeted therapy and cytotoxic therapy. Post-surgical adjuvant treatment was allowed. The value of pCR as a surrogate endpoint for clinical outcome was assessed by the means of "trial-level surrogacy", defined as the association between treatment effects on the surrogate outcome (OR of pCR between experimental and control arms) and treatment effects on the clinical outcome (HR of OS or DFS between experimental and control arms). The coefficient of determination (R2), obtained through weighted meta-regression, with 95% CIs (estimated by bootstrap methods) served as a quantificator of the surrogacy level of pCR. The analysis demonstrated a weak association between pCR and DFS [R2 = 0.08 (95% CI: 0 to 0.47; $p = 0.12$)] or OS [R2 = 0.09 (95% CI: 0.01 to 0.41; $p = 0.11$)], indicating that 8% and 9% of the variability among treatment effects on DFS and OS, respectively, was explained by the treatment effects observed on pCR. According to the authors those results supported pCR as a prognostic marker but not as a valid surrogate for long-term clinical outcomes. However, further preplanned interaction analyses showed an improved correlation between the treatment effects on pCR and the treatment effects on DFS [R2 = 0.79 (95% CI: 0.26 to 0.95); $p = 0.003$] and OS [R2 = 0.57 (95% CI: 0.19 to 0.93); $p = 0.031$] for the trials comparing intensified/dose-dense versus standard-dose chemotherapy regimens which brought authors' conclusions that "pCR may potentially meet the criteria of surrogacy with specific systemic therapies". Of note, the analysis in a subset of trials comparing chemotherapy and anti-HER2 treatment (trastuzumab) versus chemotherapy alone was not run because of insufficient number of studies (n = 2) and there was no separate analysis for HER2-positive subgroup.

The CTneoBC group [16] pooled 12 clinical trials including at least 200 primary BC patients who received preoperative chemotherapy with a median follow-up of at least 3 years. Pooled population comprised 11 955 patients, including 1 989 (17%) with HER2-positive tumours. The assessment of pCR value in the CTneoBC analysis was two-fold: (1) in a patient-level analysis EFS and OS were compared between responder and non-responder subgroups, irrespective of treatment assignment, by the estimation of hazard ratio (HR) with 95% CI and (2) in a trial-level analysis the potential of pCR as a surrogate endpoint was explored, by the assessment of the correlation between treatment effect on pCR (OR of a favourable outcome) and EFS and OS (HR of an unfavourable outcome) within a weighted regression model and estimation of R2 coefficient with 95% CI. While all included primary studies were pooled in the patient level analysis, the trial-level analysis was restricted to randomised trials. The patient-level analysis showed a significant association between pCR, defined as absence of cancer, invasive or invasive and in situ, in the breast and axillary nodes (ypT0 ypN0 or ypT0/is ypN0), with both EFS [HR = 0.44 (95% CI: 0.39 to 0.51) for ypT0 ypN0 and HR = 0.48 (95% CI: 0.43 to 0.54) for ypT0/is ypN0] and OS [HR = 0.36 (95% CI: 0.30 to 0.44) for ypT0 ypN0 and HR = 0.36 (95% CI: 0.31 to 0.42) for ypT0/is ypN0]. The relationship with clinical outcomes was weaker for pCR defined as absence of invasive cancer in the breast irrespective of ductal carcinoma in situ or nodal involvement (ypT0/is). For this reason the authors used the ypT0/is ypN0 definition for all subsequent analyses. In subgroup analyses the association between pCR and clinical outcomes was strongest in patients with most aggressive molecular subtypes: triple-negative [HR = 0.24 (95% CI: 0.18 to 0.33) for EFS and HR = 0.16 (95% CI: 0.11 to 0.25) for OS] and HER2-positive hormone receptor-negative BC [HR = 0.15 (95% CI: 0.09 to 0.27) for EFS and HR = 0.08 (95% CI: 0.03 to 0.22)

for OS]. Pathologic complete response was also associated with clinical outcomes in whole HER2-positive subgroup, irrespective of hormone receptor status [HR = 0.39 (95% CI: 0.31 to 0.50) for EFS; HR = 0.34 (95% CI: 0,24; 0,47) for OS]. In the trial-level analysis, relationship between pCR and long-term outcome was weak, for both EFS [R^2 = 0.03 (95% CI: 0.00 to 0.25)] and OS [R^2 = 0.24 (95% CI: 0.00 to 0.70)] in a pooled population of RCTs. A correlation between improvement in pCR and treatment effect on PFS and OS at the trial level was not found in subgroup analyses either, including HER2-positive subtype (R^2 not shown). The authors concluded that patients who attained pCR defined as an absence of cancer in the breast and axillary nodes (ypT0 ypN0 or ypT0/is ypN0) had improved survival and the prognostic value was greatest in aggressive tumour subtypes. At the same time they could have not validated pCR as a surrogate endpoint for improved EFS and OS.

The most recent of the included publications[19] was designed as an update and an extension of the CTneoBC study, with restriction to HER2-positive tumour subgroup. Thirty six primary studies were included, with a total population of 5768 patients, all with HER2-positive BC. Both clinical trials and cohort studies were eligible, including most recent studies (up to December 2014), published after the date of the CTneoBC systematic search (up to August 2011). As a result, the analysis of Broglio et al. included larger population of HER2-positive patients than the CTneoBC sample (3779 patients more). Similarly to the CTneoBC analysis, both patient-level (responders versus non-responders) and trial-level association between pCR and clinical endpoints was studied. The method of the trial-based analysis was comparable to the one adopted by the CTneoBC. The patient-level analysis was conducted in a different manner, with the use of a Bayesian hierarchical model, avoiding a simple pooling of results across studies, and thus incorporating variability due to between-study differences. The patient-level results were regarded as the main outcome and a basis for the conclusions, while the trial-level association was assessed mainly for a comparison with the CTneoBC results. The patient-level analysis showed that the attainment of pCR was associated both with improved EFS [median HR = 0.37 (95% probability interval: 0.32 to 0.43)] and OS [median HR = 0.34 (95% probability interval: 0.26 to 0.42)]. The effect of pCR was greater in hormone-negative patients compared with hormone-receptor positive ones and greater for neoadjuvant anti-HER2 therapy than for no preoperative anti-HER2 treatment, while there was no substantial difference between RCTs and cohort studies. In the trial-level analysis, including only RCTs and based on the weighted linear regression model, the R2 was 0.23 for EFS and 0 for OS. The correlation was stronger when the intercept of the weighted linear regression model was fixed such that a pCR OR of 1.00 corresponded to a survival HR of 1.00. In this variant the R2 was 0.63 and 0.29, respectively for EFS and OS. In addition, Broglio et al. presented the regression curves showing the expected relationship between the absolute (not relative) improvement in pCR rate and hazard ratios for EFS and OS, deemed more relevant for designing trials. Those plots showed that the accuracy of clinical outcome prediction (particularly EFS) was positively related to the observed absolute effect size on pCR: among the trials that detected an improvement in pCR rate (NOAH and NeoALTTO studies), observed HRs for EFS were in line with predicted, while no or low absolute treatment effect on pCR did not allow for accurate prediction of treatment effect on EFS. Due to lower event rate, the regression model for OS was less conclusive. The authors concluded that pCR in HER2-positive BC was associated with substantially longer times to recurrence and death, what was shown in total sample of included trials, as well as in RCTs. Thus, pCR might have been an earlier end point suitable to estimate longer-term therapeutic benefit in this BC subgroup.

Discussion

Meta-analyses included in our review differed in eligibility criteria and search dates and thus in number of included studies and patients observed. Apparent differences in conclusions though, seem to originate mostly from underlying differences in the concept of an accurate prediction and resulting variability in methodological approaches.

For unrestricted (all subtypes) BC population two "patient-level" analyses, based on comparison of clinical outcome in pooled responder versus non-responder subgroups regardless of original treatment assignment[16, 17] both showed clearly higher probability of favourable clinical outcome (defined as OS and EFS[16] or OS, DFS and RFS[17]) in patients with pCR. These results led to the conclusion that pCR was a prognostic indicator for defined clinical outcomes. Similarly, strong association of pCR with clinical outcomes was confirmed in two patient-level analyses conducted in HER2-positive subgroup [16, 19]. Only one of those studies could compare the effect of pCR between unselected BC patients and HER2-positive subgroup[16] and found out that the association was strongest in patients with most aggressive BC subtypes, including HER2-positive hormone receptor-negative BC. Significant association of pCR with favourable outcome was also maintained in this study for the whole HER2-positive subgroup, regardless of hormone receptor status.

Trial-level association between pCR and clinical outcomes was defined as the association between the treatment ef-

fect on pCR and treatment effect on clinical outcome in randomised trials. It was quantified as the percentage of clinical effect variability explained by pCR occurrence. Two papers provided trial-level analysis for population not selected by subtype [16, 18] and both of them showed only weak association between the treatment effects under consideration (< 10% of OS, EFS or DFS variance explained by pCR). As for HER2-positive subgroup, two trial-level analyses were available[16, 19] and again, the pCR effect vs clinical outcomes effect association in weighted meta-regression models was weak (about 20% variance explained for EFS but 0% for OS in Broglio et al.; in the CTneoBC paper R2 not shown). The change in meta-regression model made by Broglio et al.[19] led to notable shift of that estimation: about 60% of EFS variance and about 30% of OS variance explained by pCR in the model with fixed intercept. The conclusion of this study was not based on those results though, because of the limitations of the trial-level approach that had been indicated by the authors. In the discussion Broglio et al. point out the important limitation of the CTneoBC trial-level analysis related to inclusion mainly trials showing scarce treatment benefit on pCR (or no benefit at all), while it was shown that in the presence of absolute treatment effect on pCR the prediction of the treatment effect on EFS can be accurate. As a consequence, currently a value of the trial-level model in interpreting trial results and designing future trials is, in a case of pCR, limited.

In summary, the conclusions of the included analyses varied and those differences seem to be consistent with the variability in methodological assumptions. The conclusions for pCR surrogacy, in general patient population or HER2-positive subgroup, were positive if authors based them solely[17] or mainly[19] on the patient-level (responder) analysis, while limiting the concept of surrogacy strictly to the trial-based association[16, 18] led to rather negative conclusion. However, the most recent analysis (Broglio et al.)[19] showed that the results in the latter approach, at least for HER2-positive data subset, are strongly related to the actual size of the absolute effect on pCR observed in a clinical trial. As most of the included trials showed little or no differences in pCR rates between treatment groups, such a sample may not be feasible to assess pCR surrogacy exclusively within trial-level approach, ignoring unequivocally positive results of the responder analysis.

Of note, the authors of both papers with overall "negative" conclusions based on trial-level results[16, 18] pointed out the limitations of their approach, that might have changed the data interpretation, particularly with respect to pCR surrogacy in subpopulations with aggressive BC subtypes. Berruti et al. highlighted substantial heterogeneity of the BC population with respect to probability of obtaining pCR and prognosis and recalled individual clinical trials that indirectly suggested that pCR might have been a valid surrogate outcome in triple-negative and HER2-positive BC subpopulations. They also noted that their trial sample consisted mostly of chemotherapy trials, while studies on anti-HER2 agents would have been more relevant for pCR surrogacy assessment in HER2-positive patients.[18] In the discussion of the CTneoBC paper several limitations of the analysis were also noted, which could have obscured the existing association between pCR and long-term outcomes. Most of the included trials enrolled women with heterogeneous BC subtypes, while different molecular subtypes may respond differently to the same treatment or gain different absolute improvements in frequency of pCR. With respect to HER2-positive subgroup, it was noted that anti-HER2 treatments were used in only three trials and only one of them (the NOAH study) estimated treatment effect of adjunct HER2-agent (trastuzumab) to chemotherapy. The NOAH study differed from other included trials in terms of the effect size on pCR (absolute difference of 20% in NOAH versus 1-11% in other trials) and the difference in the proportion of patients achieving EFS at 5 years in this trial was as high as 13%. Moreover, it was found that the addition of the NOAH study to the trial-level analysis changed its results in a way suggesting possible correlation between frequency of pCR and long-term outcome. According to the authors, the substantial patient-level association between pCR and survival may justify the inference that a marked absolute increase in frequency of pCR produced by novel neoadjuvant therapy compared with standard therapy alone will translate into long-term improvements in EFS or OS.[16]

Reaching a definite conclusion on the most suitable approach to pCR validation is beyond the objective of our review. A shift in the results of trial-level meta-analyses seems quite plausible provided that clinical trials showing treatment effect on pCR are better represented. Until more such trials are published, testing new treatments improving pCR rate and with follow-up sufficient to gather meaningful survival data, it seems reasonable to account to a certain extent for the strong pCR association with long-term outcomes shown in patient-level analyses. Such a pragmatic approach was apparently adopted by regulatory authorities. In the guideline on the evaluation of anticancer medicinal products the EMA[9] highlights the need of an endpoint that would allow the assessment of efficacy at an earlier point in time than DFS and OS, for a benefit of the patients with high-risk eBC. On the basis of the CTneoBC meta-analysis the EMA argues that association between pCR and EFS in patients with aggressive tumour subtypes appears to be stronger compared to patients with less aggressive tumours and notes that in case of biomarker guided (targeted) therapy, the value of pCR

Table 2. Overview of the EMA and the FDA guidelines for the use of pCR as an endpoint for a marketing authorisation of new drug therapies for early-stage BC.

Preferred/recommended feature of an application	EMA guideline[9]	FDA guidance*[10]
Target population of patients	patients with high-risk early-stage BC, expressing suitable biomarkers in relation to the selected background regimen and the experimental compound (e.g. HER2 expression, hormone receptor status, BRCA status)	- patients with early-stage BC, judged to have a high risk of distant disease recurrence and mortality despite use of optimal modern local and systemic therapy (e.g., patients with high-grade tumours lacking ER, PR, and HER2 receptors) - patients can be classified as high risk for recurrence on the basis of conventional histologic features or by appropriately validated genomic measures, but in general should have a 5-year EFS of less than 75 percent - only populations with an unmet medical need patients with hormone receptor-positive tumours lacking high-risk features should not be enrolled
Candidate treatment	- added to standard neo(adjuvant) regimen (add-on treatment) - well-known mechanism of action of the experimental compound - neoadjuvant (preoperative treatment), i.e. systemic therapy given before lumpectomy or mastectomy to reduce the risk of breast cancer recurrence - strong biological and clinical rationale for a drug's activity in high-risk subtypes of breast cancer - an add-on treatment, added to standard therapy - postoperative cytotoxic therapy intended to treat residual disease found at the time of surgery should generally be avoided	
Comparator treatment	an established (neo)adjuvant regimen	standard therapy alone
Definition of pCR	ypT0/is ypN0 (absence of any residual invasive cancer on haematoxylin and eosin evaluation of the resected breast specimen and all sampled ipsilateral lymph nodes following completion of the neoadjuvant systemic therapy)	- ypT0/Tis ypN0 (the absence of residual invasive cancer on haematoxylin and eosin evaluation of the complete resected breast specimen and all sampled regional lymph nodes following completion of neoadjuvant systemic therapy) or - ypT0 ypN0 (the absence of residual invasive and in situ cancer on haematoxylin and eosin evaluation of the complete resected breast specimen and all sampled regional lymph nodes following completion of neoadjuvant systemic therapy)
Study design / other data	- a RCT, sufficiently large - large safety data base, with sufficiently long follow-up - supportive evidence of efficacy and safety of the experimental compound derived from studies in the metastatic setting (in most cases) - a RCT, intended to demonstrate superiority - an "add-on design", in which a standard adjuvant regimen is compared with the same regimen plus the investigational drug - a double-blind, placebo-controlled design, if blinding the investigators and patients is feasible in view of the toxicities of the investigational drug - pathologists interpreting surgical specimens for assessment of pCR should be blinded - an analysis in full intent-to-treat population - able to detect increases in pCR rate over available therapy that are of substantial magnitude - able to collect long-term safety data from a number of patients comparable to traditional adjuvant breast cancer trials	
Results	- a major increase in pCR - only minor add-on changes in toxicity	a clinically meaningful, not only statistically significant, difference in pCR
Post-approval obligations	confirmatory study data in terms of EFS/DFS/OS (achieved through prolonged follow-up of the neoadjuvant study if sufficiently large or through a separate adjuvant study)	- the confirmatory trial (ongoing at the time of accelerated approval) should demonstrate a clinically meaningful and statistically significant improvement in EFS, DFS, or OS - the confirmatory data may be derived from the original randomised neoadjuvant trial that supported the accelerated approval (in a longer follow up) or from a separate, larger trial in either the neoadjuvant or adjuvant setting - additional safety trials may be required

** the FDA guidance contains nonbinding recommendations, applicable to the accelerated approval procedure*
Abbreviations: AJCC - American Joint Committee on Cancer; BC - breast cancer; BRCA - breast cancer gene; DFS – disease-free survival; EFS – event-free survival; EMA - European Medicines Agency; FDA - U.S. Food and Drug Administration; HER2 – human epidermal growth factor receptor 2; pCR – pathologic complete response; OS – overall survival; RCT – randomised controlled trial

as outcome measure may be limited if the biomarker is not associated with aggressive tumours. Thus, despite the fact that "the true surrogacy" of pCR has not been established, the EMA finds an approval based on pCR acceptable if several, specified additional conditions are satisfied (Table 2).

The key conditions include the diagnosis of high-risk eBC, well-characterised mechanism of action of the novel agent and a major increase in pCR rate together with only minor changes in toxicity. Furthermore, confirmatory data should be submitted for clinical outcomes in the post-approval period. Those data can be derived from the neoadjuvant study, with prolonged follow-up or from the separate adjuvant study.[9] Likewise, the FDA accepts pCR as an endpoint supporting the accelerated approval of novel systemic therapies.[10] The accelerated approval regulations themselves address the areas of unmet medical need within populations of patients with serious or life-threatening diseases. It was recognized that a significant unmet need remains for certain high-risk or poor prognosis subgroups of eBC patients. For this reason developing highly effective new drugs for these populations was set as a priority. As a part of the rationale for use pCR as a surrogate endpoint in neoadjuvant trials FDA points out that, as a result of improvements in eBC therapy, demonstrating an adequate difference in DFS or OS requires RCTs with large sample sizes and prolonged follow-up. Consequently, time from initiation of a phase 3 trial of a drug in metastatic BC to approval for its use in eBC patients (as an adjuvant treatment) has extended to a decade or more. A pCR, in turn, may be reasonably likely to predict clinical benefit and can be assessed within several months of initiation of an investigational drug used preoperatively. In the guidance the results and limitations of the CTneoBC meta-analysis are discussed in detail. Although no correlation between magnitude of difference in pCR rates between treatment arms and EFS or OS at a trial level were found, FDA considers it to be "reasonably likely" that a novel agent that produces a marked absolute increase in pCR rate compared with standard therapy alone in the full intent-to-treat population may result in long-term improvements in EFS or OS[10] (p7). Similarly to EMA, FDA gives additional recommendations for the use of pCR as an endpoint for accelerated approval, including a preference for a randomised, add-on study design with a superiority hypothesis and a prespecified, target magnitude of expected pCR effect. The appropriate patient populations are those with a high risk of distant disease recurrence and mortality despite use of optimal modern local and systemic therapy, with an unmet medical need. For subsequent conversion to regular approval, the confirmatory trial, supposed to be ongoing at the time of accelerated approval, should demonstrate a clinically meaningful and statistically significant improvement in EFS, DFS, or OS.[10]

The National Institute for Health and Care Excellence's guidance for neoadjuvant pertuzumab for the neoadjuvant treatment of HER2-positive BC[15] may serve an example of a pragmatic approach to uncertainty related to pCR-based technology appraisal on the level of national reimbursement decisions. The key requirement for positive NICE recommendation for a new drug technology is a reliable demonstration of cost-effectiveness, that is, the cost of quality-adjusted life-year (QALY) gained with the new therapy should, in general, not exceed £20,000- £30,000 threshold.[20] In case of neoadjuvant pertuzumab, due to the absence of reliable estimate of EFS or OS gain from the randomised study (NeoSphere), long-term clinical benefit had to be modelled on the basis of demonstrated pCR improvement. Consequently, the validity of clinical outcome prediction on the basis of treatment effect of pCR was a key condition for the economic submission to be accepted as reliable. The appraisal committee considered studies that investigated the value of pCR as a clinically meaningful indicator of EFS and OS, with particular emphasis on the CTneoBC meta-analysis, and agreed that there was considerable uncertainty whether pCR could be viewed as a surrogate marker of long-term benefit. Yet it was also acknowledged that according to the clinical experts, obtaining a response can have an important psychological benefit for a patient, that a pCR indicates that not only tumour cells but also any micro-metastases are likely to have been treated and that a reduction or disappearance of tumour in the breast potentially allows for less radical surgery. Furthermore, the committee found it more likely that the complete disappearance of cancer in the breast and nodes is associated with improved long-term outcomes than that there is no such relation. The existing uncertainty in cost-effectiveness estimation was dealt with choosing conservative assumptions in the economic model and additional financial agreements with MAH. As a result NICE has recommended pertuzumab use, within its marketing authorisation: for the neoadjuvant treatment of adults with HER2-positive breast cancer, that is, in patients with HER2-positive, locally advanced, inflammatory or early-stage breast cancer at high risk of recurrence.[15]

The strength of our work is a presentation of the most comprehensive and generalisable meta-analyses, deemed most reliable due to systematic selection of primary data sources on the basis of clearly defined eligibility criteria. We aimed to systematically review the methodology and results of the included meta-analyses, to improve the understanding of possible sources of existing controversies and variability in conclusions on a pCR value as an surrogate endpoint in neoadjuvant BC trials. The meta-analytic results were put into the context of current regulatory and reimbursement decision-mak-

ing. As a consequence of adopted criteria for inclusion, several published meta-analyses were not eligible for our review, including pooled analyses of German neoadjuvant trials[21] and of prospective neoadjuvant studies of the Japanese Breast Cancer Society.[22, 23] As in the former the analysis of predictive value of pCR was focused on patients with infiltrating lobular BC and the latter were restricted to Japanese patients, we believe that the exclusion of those studies did not bias our conclusions. The sample of studies included in our review was sufficient to show a diversity of methodological approaches to the subject and the way in which those differences may underlie existing discrepancies in the data interpretation by authors of the analyses themselves and by decision-makers at the stage of regulatory approval, as well as at the level of subsequent national reimbursement decisions.

Conclusions

Currently the use of pathologic complete response as a surrogate of long-term clinical outcomes in clinical trials of new systemic therapies for early-stage breast cancer is not unambiguously supported by meta-analyses. It was recognised, however, that significant heterogeneity of breast cancer population and low number of studies showing substantial treatment effect on pCR could bias the results of "trial-level" analyses. In contrast to the weakness of trial-level correlation, pooled studies consistently show significant association of pCR with clinical outcomes on the individual patient-level, particularly strong for aggressive tumour subtypes – including HER2-positive breast cancer. That inconsistency is the source of uncertainty, which leads to conflicting data interpretation by researchers and health-care decision-makers. On the one hand the assessment of new drug therapies solely on the basis of pCR would carry the risk that not effective treatments will be used and reimbursed. On the other hand the urgent need for new therapies in patient subgroups with poor prognosis still remains and is difficult to meet without the possibility of early assessment and approval. The existing EMA and FDA guidelines show how to make use of the existing evidence, in spite of its limitations, in a pursuit of satisfying the unmet medical needs. Several ways of mitigating the risk of suboptimal decisions have been proposed, including restriction of the early, pCR-based approval to add-on therapies for high-risk patients and requirement of further, post-approval data. We believe that unmet needs related to high-risk breast cancer patient populations should not be overlooked also by national HTA bodies, whose negative decisions may block access to already approved therapies. The example of NICE guideline for neoadjuvant pertuzumab in HER2-positive breast cancer shows, that in case of pCR-approved therapies the uncertainty related to cost-effectiveness can be reduced with adopting conservative approach in economic model and undertaking additional financial/risk-sharing agreements.

Acknowledgements

This study was supported by a grant from Roche Polska Sp. z o.o., Warsaw, Poland.

References

1. European network for Health Technology Assessment (EUnetHTA). Guideline. Endpoints used for Relative Effectiveness Assessment: Clinical Endpoints. Adapted version (2015) based on "Endpoints used for Relative Effectiveness Assessment of pharmaceuticals: Clinical Endpoints" - February 2013. Available from: http://www.eunethta.eu/sites/default/files/WP7-SG3-GL-clin_endpoints_amend2015.pdf [Accessed November 3, 2017].
2. Jönsson B. Technology assessment for new oncology drugs. Clin Cancer Res. 2013;19(1):6-11.
3. Jönsson B. Bringing in health technology assessment and cost-effectiveness considerations at an early stage of drug development. Mendelsohn J, Ringborg U, Schilsky R, eds. Molecular Oncology. 2015;9(5):1025-1033.
4. International Agency for Research on Cancer. Globocan 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. Cancer fact sheets. Breast Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx [Accessed November 6, 2017].
5. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al.; Cancer Intervention and Surveillance Modeling Network (CISNET) Collaborators. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 2005;353(17):1784-92.
6. Puliti D, Zappa M. Breast cancer screening: are we seeing the benefit? BMC Medicine 2012;10:106.
7. Malhotra GK, Zhao X, Band H, Band V. Histological, molecular and functional subtypes of breast cancers. Cancer Biol Ther. 2010;10(10):955-60.
8. Jhan JR, Andrechek ER. Effective personalized therapy for breast cancer based on predictions of cell signaling pathway activation from gene expression analysis. Oncogene 2017;36(25):3553-3561.
9. European Medicines Agency (EMA), Committee for Medicinal Products for Human Use (CHMP).

Appendix 4 to the guideline on the evaluation of anticancer medicinal products in man. 17 December 2015, EMA/CHMP/703715/2012 Rev. 2. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_001124.jsp&mid=WC0b01ac0580034cf3 [Accessed November 3, 2017].

10. U.S. Department of Health and Human Services Food and Drug Administration (FDA) Center for Drug Evaluation and Research (CDER). Guidance for Industry Pathological Complete Response in Neoadjuvant Treatment of High-Risk Early-Stage Breast Cancer: Use as an Endpoint to Support Accelerated Approval. October 2014. Available from: https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064981.htm [Accessed November 3, 2017].
11. Martinalbo J, Bowen D, Camarero J, Chapelin M, Démolis P, Foggi P, et al. Early market access of cancer drugs in the EU. Ann Oncol. 2016;27(1):96-105.
12. Pan-Canadian Oncology Drug Review (pCODR). PCODR Expert Review Committee (pERC) Final Recommendation. Final Recommendation for Pertuzumab (Perjeta) for Neoadjuvant Breast Cancer pERC Meeting: April 17, 2015; pERC Reconsideration Meeting: July 2, 2015. Available from: https://www.cadth.ca/perjeta-or-perjeta-herceptin-combo-pack-neoadjuvant-breast-cancer-details [Accessed November 6, 2017].
13. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWIG). IQWiG Reports – Commission No. A15-34. Pertuzumab (new therapeutic indication) – Benefit assessment according to §35a Social Code Book V. Extract of dossier assessment A15-34, Version 1.0, 27 November 2015.x. Available from: https://www.iqwig.de/en/projects-results/projects/drug-assessment/a15-34-pertuzumab-new-therapeutic-indication-benefit-assessment-according-to-35a-social-code-book-v-dossier-assessment.6943.html [Accessed November 6, 2017].
14. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWIG). IQWiG Reports – Commission No. A16-01. Pertuzumab – Addendum to Commission A15-34. Version 1.0, 29 January 2016. Available from: https://www.iqwig.de/en/projects-results/projects/drug-assessment/a16-01-pertuzumab-addendum-to-commission-a15-34.7183.html [Accessed November 6, 2017].
15. National Institute for Health and Care Excellence (NICE). Pertuzumab for the neoadjuvant treatment of HER2-positive breast cancer (TA424). Technology appraisal guidance Published: 21 December 2016. Available from: https://www.nice.org.uk/guidance/ta424 [Accessed November 6, 2017].
16. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 2014;384(9938):164-72.
17. Kong X, Moran MS, Zhang N, Haffty B, Yang Q. Meta-analysis confirms achieving pathological complete response after neoadjuvant chemotherapy predicts favourable prognosis for breast cancer patients. Eur J Cancer 2011;47(14):2084-90.
18. Berruti A, Amoroso V, Gallo F, Bertaglia V, Simoncini E, Pedersini R, et al. Pathologic complete response as a potential surrogate for the clinical outcome in patients with breast cancer after neoadjuvant therapy: a meta-regression of 29 randomized prospective studies. JClin Oncol. 2014;32(34):3883-91.
19. Broglio KR, Quintana M, Foster M, Olinger M, McGlothlin A, Berry SM, et al. Association of Pathologic Complete Response to Neoadjuvant Therapy in HER2-Positive Breast Cancer With Long-Term Outcomes: A Meta-Analysis. JAMA Oncol. 2016;2(6):751-60.
20. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013 (PMG9). Published: 4 April 2013. Available from: https://nice.org.uk/process/pmg9 [Accessed November 9, 2017].
21. Loibl S, Volz C, Mau C, Blohmer JU, Costa SD, Eidtmann H, et al. Response and prognosis after neoadjuvant chemotherapy in 1,051 patients with infiltrating lobular breast carcinoma. Breast Cancer Res Treat. 2014;144(1):153-62.
22. Kuroi K, Toi M, Ohno S, Nakamura S, Iwata H, Masuda N, et al. Prognostic significance of subtype and pathologic response in operable breast cancer; a pooled analysis of prospective neoadjuvant studies of JBCRG. Breast Cancer. 2015;22(5):486-95.
23. Kuroi K, Toi M, Ohno S, Nakamura S, Iwata H, Masuda N, et al. Comparison of different definitions of pathologic complete response in operable breast cancer: a pooled analysis of three prospective neoadjuvant studies of JBCRG. Breast Cancer 2015;22(6):586-95.