Disease-free survival as a surrogate for overall survival in patients with HER2-positive, early breast cancer in trials of adjuvant trastuzumab for up to 1 year: a systematic review and meta-analysis

Everardo D Saad, Pierre Squifflet, Tomasz Burzykowski, Emmanuel Quinaux, Suzette Delaloge, Dimitris Mavroudis, Edith Perez, Martine Piccart-Gebhart, Bryan P Schneider, Dennis Slamon, Norman Wolmark, Marc Buyse

Summary

Background Although frequently used as a primary endpoint, disease-free survival has not been validated as a surrogate for overall survival in early breast cancer. We investigated this surrogacy in the adjuvant setting of treatment with anti-HER2 antibodies.

Methods In a systematic review and meta-analysis, we identified published and non-published randomised controlled trials with completed accrual and available disease-free survival and overall survival results for the intention-to-treat population as of September 2016. Bibliographic databases (MEDLINE, Embase, and Cochrane Central Register of Controlled Trials), clinical trial registries (Clinicaltrials.gov, EU Clinical Trials Register, WHO International Clinical Trials Registry Platform, and PharmNet.Bund), and trial registries from relevant pharmaceutical companies were searched. Eligibility for treatment of HER2-positive early breast cancer required at least one group to have an anti-HER antibody treatment (ie, trastuzumab, pertuzumab, or trastuzumab emtansine) planned for 12 months, and at least one control arm with chemotherapy without the antibody, a lower total dose or duration of the antibody, or observation alone. Units of analysis were contrasts: two-group trials gave rise to one contrast, whereas trials with more than two groups gave rise to more than one contrast. We excluded trials enrolling patients with recurrent, metastatic, or non-invasive disease, and those testing neoadjuvant therapy exclusively. Our primary objective was to estimate patient-level and trial-level correlations between disease-free survival and overall survival. We measured the association between disease-free survival and overall survival using Spearman's correlation coefficient (r_s), and the association between hazard ratios (HRs) for disease-free survival and overall survival using R^2. We computed the surrogate threshold effect, the maximum HR for disease-free survival that statistically predicts an HR for overall survival less than 1·00 in a future trial.

Findings Eight trials (n=21 480 patients) gave rise to a full set (12 contrasts). Patient-level associations between disease-free and overall survival were strong (r_s=0·90 [95% CI 0·89–0·90]). Trial-level associations gave rise to values of R^2 of 0·75 (95% CI 0·50–1·00) for the full set. Subgroups defined by nodal status and hormone receptor status yielded qualitatively similar results. Depending on the expected number of deaths in a future trial, the surrogate threshold effects ranged from 0·56 to 0·81, based on the full set.

Interpretation These findings suggest that it is appropriate to continue to use disease-free survival as a surrogate for overall survival in trials in HER-2-positive, early breast cancer. The key limitation of this study is the dependence of its results on the trials included and on the existence of an outlying trial.

Funding Roche Pharma AG.

Lancet Oncol 2019; 20: 361–70

Published Online
January 29, 2019
http://dx.doi.org/10.1016/S1470-2045(18)30750-2

This online publication has been corrected. The corrected version first appeared at thelancet.com/oncology on March 1, 2019

See **Comment** page 315

International Drug Development Institute, Louvain-la-Neuve, Belgium (E D Saad MD, P Squifflet MSc, Prof T Burzykowski PhD, E Quinaux MSc); **Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium** (Prof T Burzykowski, M Buyse ScD); **Unicancer Breast Group, Paris, France** (S Delaloge MD); **University of Crete, Heraklion, Greece** (Prof D Mavroudis MD); **Mayo Clinic, Jacksonville, FL, USA** (Prof E Perez MD); **Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium** (Prof M Piccart-Gebhart PhD); **Indiana University, Indianapolis, IN, USA** (B P Schneider MD); **University of California, Los Angeles (UCLA), Los Angeles, CA, USA** (Prof D Slamon PhD); **NRG Oncology/Pittsburgh, Pittsburgh, PA, USA** (Prof N Wolmark MD); **and International Drug Development Institute, San Francisco, CA, USA** (M Buyse)

Correspondence to:
Dr Everardo D Saad, International Drug Development Institute, 1340 Louvain-la-Neuve, Belgium
everardo.saad@iddi.com

Introduction

Breast cancer is divided into molecular subtypes with relevant prognostic and predictive implications for clinical practice.[1] In HER2-positive breast cancer, adjuvant therapy with trastuzumab improves outcomes.[2–4] Although shorter and longer durations of trastuzumab therapy,[2,5–8] as well as other anti-HER2 agents,[9–11] have been investigated in the adjuvant setting, 1 year of trastuzumab remains the most common option for treatment.[1,12] The development of novel adjuvant regimens is a lengthy process, and the analysis of overall survival requires a long follow-up. One possibility to expedite drug development and patient access to improved regimens is to test new drugs as neoadjuvant therapies, because drugs that perform well in this context are more promising in the adjuvant setting than regimens that add no improvements in clinical or pathological responses when used as neoadjuvant treatment. However, doubts remain about the predictive ability of the neoadjuvant platform,[13,14] and another

Research in context

Evidence before this study

Although disease-free survival has often been used as a primary endpoint in adjuvant trials of breast cancer, it has not been formally validated as a surrogate for overall survival.

This systematic review and meta-analysis was done to evaluate the role of disease-free survival as a surrogate for overall survival in the adjuvant treatment of HER2-positive early breast cancer. In October, 2015, and October, 2016, bibliography databases (MEDLINE, Embase, and Cochrane Central Register of Controlled Trials), clinical trial registries (Clinicaltrials.gov, EU Clinical Trials Register, WHO International Clinical Trials Registry Platform, and PharmNet.Bund), and trial registries from relevant pharmaceutical companies were searched for published or unpublished randomised clinical trials in HER2-positive early breast cancer on anti-HER2 monoclonal antibodies. Search terms were "breast cancer" and synonyms; "trastuzumab", "pertuzumab", or "trastuzumab emtansine", and synonyms; and "phase 2" or "phase 3" or "phase 4". Data were provided by investigators for all but one of the nine eligible trials.

Added value of this study

A two-level modelling approach was used to estimate the association between disease-free survival and overall survival (patient-level association) and between the treatment effects on these endpoints (trial-level association). Patient-level associations were strong (r_s=0·90 [95% CI 0·89–0·90]). Trial-level associations were moderate or strong, depending on the set analysed (R^2 0·75 [95% CI 0·50–1·00] in the full analysis set, 0·84 [0·67–1·00] in the reduced set). These results suggest that disease-free survival has good overall statistical validity as a surrogate for overall survival in the adjuvant treatment of HER2-positive, early breast cancer. These results apply mainly to the adjuvant use of trastuzumab for 12 months and cannot be readily extrapolated to other types of adjuvant anti-HER2 therapy, whether changing the agent (trastuzumab) or the duration of treatment.

Implications of all the available evidence

The levels of association could be sufficient for the purpose of replacing a final endpoint such as overall survival. This surrogacy has implications for trial design as well as for the approval of novel agents. Further studies should assess disease-free survival as a surrogate for overall survival in other breast cancer phenotypes.

possibility to expedite the development of adjuvant therapy is to use surrogates for overall survival.[15] Although disease-free survival has often been used as the primary endpoint in adjuvant trials of breast cancer, to our knowledge it has not been formally validated in this setting, as it has in others—for example, colon, stomach, and lung cancer.[16–18] Meta-analyses of individual-patient data from randomised trials provide two measures of association between the potential surrogate and the final endpoint of interest: the patient-level associations, and the trial-level associations.[19] Patient-level associations denote the prognostic role of the surrogate (eg, whether patients with prolonged disease-free survival are also more likely to have prolonged overall survival), whereas trial-level associations provide predictive information (ie, whether treatment-induced changes in the surrogate endpoint are accompanied by proportional changes in the final endpoint). These two associations are independent, as shown by a strong patient-level association but weak trial-level association between pathological complete response and overall survival in a pooled analysis of neoadjuvant therapy trials in breast cancer.[14] Our study was done to evaluate the role of disease-free survival as a surrogate for overall survival in the adjuvant treatment of HER2-positive breast cancer.

See Online for appendix

Methods

Search strategy and selection criteria

In this systematic review and meta-analysis, eligible studies were randomised clinical trials of adjuvant therapy for patients with stage I–III breast cancer, with randomisation done after surgery and accrual completed as of September, 2016; patients had to have HER2-positive disease, either exclusively or with stratification for HER2 positivity ascertained by the accepted methods of immunohistochemistry and fluorescent in-situ hybridisation (of note, patients with HER2-negative disease were not eligible for the trials analysed). Anti-HER2 antibody use had to have been planned for a total duration of 1 year in at least one of the trial groups, and at least one group had to receive observation alone, chemotherapy alone, or a lower total dose (per cycle or in terms of treatment duration) of the anti-HER2 antibody. Finally, the trial research question must have involved the anti-HER2 antibody. We excluded trials that enrolled patients with recurrent, metastatic, or non-invasive disease, and those that tested neoadjuvant therapy exclusively (if both neoadjuvant and adjuvant therapies were allowed in a trial, randomisation must have been after surgery). The full list of selection criteria is provided in the appendix (p 3).

The initial search for eligible trials was done by a third party, HealthEcon (Basel, Switzerland), in October, 2015, and updated in October, 2016. Bibliography databases (MEDLINE, Embase, and Cochrane Central Register of Controlled Trials), clinical trial registries (Clinicaltrials.gov, EU Clinical Trials Register, WHO International Clinical Trials Registry Platform, and PharmNet.Bund), and trial registries from relevant pharmaceutical companies were searched for published or unpublished

randomised clinical trials in HER2-positive early breast cancer on the anti-HER2 monoclonal antibodies trastuzumab, pertuzumab, or trastuzumab emtansine. No language restrictions were applied. Search terms were "breast cancer" and synonyms, "trastuzumab", "pertuzumab", or "trastuzumab emtansine", and synonyms, and "phase 2", "phase 3", or "phase 4". The full list of search terms and further details are available on request, and the MEDLINE search done on Oct 9, 2015, is shown in the appendix (pp 4–5). The title and abstract of the trials and publications were screened for relevance on the basis of the prespecified selection criteria. Trials and publications excluded as not relevant at this stage were not further documented. Trials and publications with potential relevance after the first stage of screening were included in full-text screening. Trials and publications excluded after this stage were documented, with their reason for exclusion. The search and selection were done independently by two individuals mentioned in the Acknowledgments. If there was debate concerning the inclusion at the full-text screening stage, this was resolved by consensus.

The study protocol (appendix pp 10–24) was approved by the Ethics Committee of Hasselt University, Diepenbeek, Belgium. Because the study consisted of re-analysis of data from clinical trials that were approved by institutional review boards, no informed consent was sought from patients.

Data extraction

The only data used in the statistical analysis extracted directly from publications were the accrual dates. Data items requested from investigators are shown in the protocol (appendix p 16). Data maturity for disease-free survival and overall survival was arbitrarily defined, with immaturity meaning that results were not yet published at the time the analysis was done. Investigators of potentially eligible trials were contacted and asked if individual-patient data for disease-free survival and overall survival were available and could be shared by June 30, 2017.

Data from one eligible trial that finished accrual in 2010 (the HARE trial[6]) were not made available because of concerns on the part of the French National Cancer Institute about whether repeated informed consent would need to be obtained from patients. Two eligible trials were identified for which data would not be available in a timely fashion, but whose results have been presented when our study was ongoing.[20,21] The results from these two trials and those from the PHARE trial[6] were used in an exploratory analysis post-hoc to verify some of the predictions resulting from this study. Moreover, a third eligible trial was ongoing and had its results presented in a scientific meeting after the current analyses had been finalised. Data from this third trial were not used in this study.[22] No additional eligible trials with available data have been identified by us. One further trial published only in abstract form was found incidentally before the analysis, and in this case, contact could not be established with investigators; this trial, done in India and not reported in any of the registries accessed was found in the proceedings of the European Society of Medical Oncology meetings for another reason, enrolled 134 patients and only reported 3-year disease-free survival results.[23]

Outcomes

The primary objective of the study was to assess surrogacy of disease-free survival for overall survival in trials that used an adjuvant anti-HER2 antibody, considering the intention-to-treat population of each trial. Secondary objectives were to estimate these associations in trials that had given trastuzumab for 1 year in at least one experimental group, and at least one non-anti-HER2-containing group, with the same chemotherapy in both of these groups; to estimate these associations in patient subgroups defined by hormone receptor expression and nodal status; to do sensitivity analyses according to previous use of neoadjuvant therapy and different definitions of disease-free survival;[24] and to do exploratory

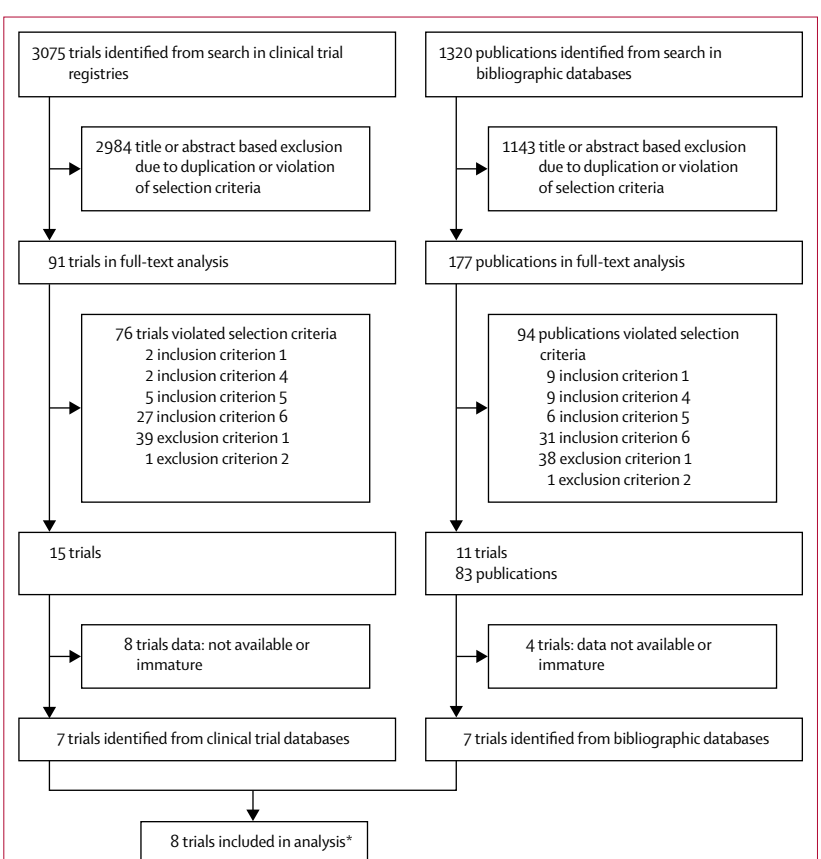

Figure 1: **Study selection**
Details of inclusion and exclusion criteria are provided in the appendix p 3. Data maturity for disease-free survival and overall survival was arbitrarily defined, with immaturity meaning that results were not yet published at the time the analysis was done. *See the Results for description of how these 8 trials were arrived at.

	Eligible patients	Eligible arms (contrasts)	Events defining disease-free survival	Control arm(s) for this review	Experimental arm(s) for this review
ALTTO[9] (NCT00490139)	8381	4 (2)	Recurrence of invasive breast cancer at local, regional, or distant sites; contralateral invasive breast cancer; second non-breast malignancy; or death from any cause	Chemotherapy* plus lapatinib for 12 months; chemotherapy* plus trastuzumab for 12 months	Chemotherapy* plus lapatininb and trastuzumab for 12 months; chemotherapy* plus trastuzumab for 12 weeks and lapatinib for 34 weeks
BCIRG 006[4] (NCT00021255)	3222	3 (2)	Breast cancer recurrence, a second primary cancer (excluding contralateral ductal carcinoma in situ), or death from any cause	Doxorubicin, cyclophosphamide, and docetaxel	Doxorubicin, cyclophosphamide, docetaxel, and trastuzumab for 12 months; docetaxel, carboplatin, and trastuzumab for 12 months
E2198[26] (NCT00003992)	234	2 (1)	Disease recurrence, development of invasive second primary, or death	Paclitaxel and trastuzumab for 12 weeks followed by doxorubicin and cyclophosphamide for four cycles	Same as control, plus trastuzumab for 12 months
HERA[2] (NCT00045032)	3401	2 (1)	Recurrence of breast cancer at any site, development of ipsilateral or contralateral breast cancer (including ductal carcinoma in situ but not lobular carcinoma in situ), second non-breast malignancy (other than basal-cell or squamous-cell carcinoma of the skin or carcinoma in situ of the cervix), or death from any cause	Observation	Trastuzumab for 12 months
HORG[7] (NCT00615602)†	481	2 (1)	Breast cancer recurrence (either locoregional or distant), contralateral breast cancer, second non-breast malignancy, or death from any cause	Dose-dense fluorouracil, epirubicin, and cyclophosphamide followed by dose-dense docetaxel plus trastuzumab for 6 months	Dose-dense fluorouracil, epirubicin and cyclophosphamide followed by dose-dense docetaxel plus trastuzumab for 12 months
NCCTG N9831[3,25] (NCT00005970)	1885	3 (3)	Local, regional, or distant recurrence; contralateral breast cancer, including ductal carcinoma in situ; other second primary cancers; or death before recurrence or a second primary cancer	Doxorubicin and cyclophosphamide followed by weekly paclitaxel	Same as control plus sequential trastuzumab for 12 months while group C‡ was open; same as control plus concurrent trastuzumab for 12 months; same as control plus sequential trastuzumab for 12 months while group C‡ was closed
NSABP B-31[3] (NCT00004067)	2102	2 (1)	Local, regional, or distant recurrence; contralateral breast cancer, including ductal carcinoma in situ; other second primary cancers; or death before recurrence or a second primary cancer	Doxorubicin and cyclophosphamide followed by paclitaxel every 3 weeks	Same as control, plus trastuzumab for 12 months
PACS 04[27] (NCT00054587)	527	2 (1)	Local or regional recurrence, distant metastases, contralateral breast cancer, or death from any cause	Fluorouracil, epirubicin, and cyclophosphamide, or epirubicin and docetaxel	Same as control, plus trastuzumab for 12 months

*Chemotherapy of choice prior to randomisation (design 1), anthracycline-based regimen followed by taxane and concurrent trial therapy (design 2A), or docetaxel, carboplatin, and concurrent trial therapy (design 2B). †Because of methodological issues with this study, it was excluded from some analyses (post-hoc reduced set analyses). ‡This trial had a period during which accrual was closed to one of the groups (C, with concurrent chemotherapy and trastuzumab), thus generating the need to create contrasts that only had concurrently randomised patients.

Table 1: **Selected characteristics of trials analysed and designation of control and experimental groups for this systematic review**

analyses of the association between disease-free survival and breast cancer-specific survival. Because no information was generally available on causes of death, in this exploratory analysis we considered deaths preceded by recurrence to be due to breast cancer.

The main variables of interest within each eligible trial were age, tumour stage, lymph node stage, hormone receptor (both oestrogen and progesterone) status, and dates of randomisation, disease relapse, death, or latest follow-up. Because access to all data from each trial was granted by the respective investigators, a uniform method could be used to verify the adequacy of the data for the planned analyses—namely, data cleaning and management performed by the statistician.

Statistical analysis

A two-level modelling approach was used to estimate the association between disease-free survival and overall survival and between the treatment effects on these endpoints.[19] At the patient level, the joint distribution of the surrogate endpoint and the true endpoint was estimated with a copula-based model. Three different distributions, corresponding to three different copulas (Clayton's, Hougaard's, or Plackett's) were considered, and the one providing the maximum likelihood value was selected for inference. Subsequently, the strength of the association between the surrogate and the true endpoint was quantified by the value of Spearman's rank correlation coefficient (r_s) corresponding to the selected copula. For the trial-level assessment, the proportional hazards model was used to jointly estimate the hazard ratios (HRs) for disease-free survival and overall survival. A linear regression was then fitted through the points representing the logarithms of the hazard ratio (logHR) for disease-free survival and for overall survival from each unit of analysis, which was termed contrast. This term was used to make a distinction between the comparison of treatment groups

	Contrast	Number of patients	Disease-free survival			Overall survival		
			Median follow-up (months)	Number of events overall	HR (95% CI)	Median follow-up (months)	Number of events overall	HR (95% CI)
ALTTO[9]	Chemotherapy plus lapatinib for 12 months *vs* chemotherapy plus lapatinib plus trastuzumab for 12 months	4193	82 (72–85)	776	0·68 (0·59–0·78)	83 (72–86)	396	0·69 (0·56–0·84)
ALTTO[9]	Chemotherapy plus sequential trastuzumab and lapatinib *vs* chemotherapy plus trastuzumab for 12 months	4188	83 (72–85)	730	1·08 (0·93–1·25)	83 (72–86)	362	1·16 (0·94–1·43)
BCIRG 006[4]	Doxorubicin, cyclophosphamide, and docetaxel *vs* same chemotherapy plus trastuzumab for 12 months	1611	126 (105–128)	419	0·68 (0·56–0·83)	126 (103–128)	247	0·62 (0·48–0·79)
BCIRG 006[4]	Doxorubicin, cyclophosphamide, and docetaxel *vs* docetaxel, carboplatin, and trastuzumab for 12 months	1611	126 (107–128)	424	0·81 (0·66–0·98)	126 (104–128)	264	0·81 (0·63–1·03)
E2198[26]	Paclitaxel and trastuzumab for 12 weeks followed by doxorubicin and cyclophosphamide *vs* same plus trastuzumab for 12 months	234	76 (71–80)	62	1·25 (0·76–2·07)	77 (73–83)	42	1·25 (0·68–2·29)
HERA[2]	Observation *vs* trastuzumab for 12 months	3401	132 (121–139)	1113	0·76 (0·68–0·86)	132 (121–138)	725	0·74 (0·64–0·86)
HORG[7]	Fluorouracil, epirubicin, cyclophosphamide, and docetaxel plus trastuzumab for 6 months *vs* fluorouracil, epirubicin, cyclophosphamide, and docetaxel plus trastuzumab for 12 months	481	49 (33–70)	45	0·63 (0·35–1·16)	50 (33–70)	18	1·45 (0·57–3·67)
NCCTG N9831[3,25]	Doxorubicin, cyclophosphamide, and paclitaxel *vs* doxorubicin, cyclophosphamide and paclitaxel plus sequential trastuzumab for 12 months, group C open	1423	146 (119–162)	412	0·82 (0·67–1·00)	153 (136–169)	295	0·79 (0·62–1·00)
NCCTG N9831[3,25]	Doxorubicin, cyclophosphamide, and paclitaxel *vs* doxorubicin, cyclophosphamide and paclitaxel plus concurrent trastuzumab for 12 months	1418	146 (117–165)	399	0·70 (0·57–0·86)	153 (137–169)	259	0·75 (0·58–0·96)
NCCTG N9831[3,25]	Doxorubicin, cyclophosphamide, and paclitaxel *vs* doxorubicin, cyclophosphamide and paclitaxel plus sequential trastuzumab for 12 months, group C closed	291	165 (134–178)	105	0·56 (0·38–0·83)	175 (158–182)	75	0·73 (0·46–1·16)
NSABP B-31[3]	Doxorubicin, cyclophosphamide, and paclitaxel *vs* same plus trastuzumab for 12 months	2102	119 (101–134)	696	0·59 (0·50–0·68)	119 (102–134)	445	0·66 (0·54–0·79)
PACS 04[27]	Fluorouracil, epirubicin, and cyclophosphamide, or epirubicin, and docetaxel *vs* same plus trastuzumab for 12 months	527	113 (93–118)	190	0·76 (0·57–1·02)	112 (94–118)	106	0·81 (0·55–1·19)

HR=hazard ratio.

Table 2: **Selected results for each contrast analysed**

in the original trials and the comparisons of two treatments used in our study. Such distinction is only necessary for trials with more than two groups. For two-group trials, each contrast corresponds to the same comparison between treatments given in the original trial. Two trials with three groups[4,25] had their chemotherapy-alone group randomly split to generate two or more contrasts. One trial with four groups gave rise to two contrasts with no need for random splits.[9] Random splitting of the control group does not lead to multiplicity issues, because no extra significance tests are generated. For simplicity of interpretation, the graphs show HRs (rather than logHRs) in their axes. Given that all identified trials had trastuzumab as the anti-HER2 antibody, and to enable more homogeneous comparisons, trastuzumab for 1 year was used in the numerator for disease-free survival and overall survival for all contrasts analysed; this was done regardless of whether trastuzumab was used alone,[2] with chemotherapy,[3,4,7,25–27] or with lapatinib,[9] allowing for meaningful interpretation of predictions on the basis of the obtained regression line.

In all analyses, an attempt was made to fit the regression models while taking into account the estimation error present in the estimated HRs for overall survival and disease-free survival by use of a measurement-error model.[19] In case of numerical problems with fitting the models, we used the numbers of deaths in each contrast to produce weighted regression models. The linear regression fitted through the estimated treatment effects provides a coefficient of determination (R^2), which quantifies the proportion of variance in the effects of treatment on the true endpoint, which is explained by the surrogate. Additionally, the fitted regression line (whether obtained by use of the measurement-error modelling or weighted regression) allows construction of a 95% prediction interval for the HR for overall survival corresponding to a particular value of HR for disease-free survival; the prediction interval estimates, with 95% probability, the interval in which the HR for overall survival will fall. The 95% prediction interval based on the weighted regression model depends on the weight assigned to the contrast for the HR being predicted (usually taken

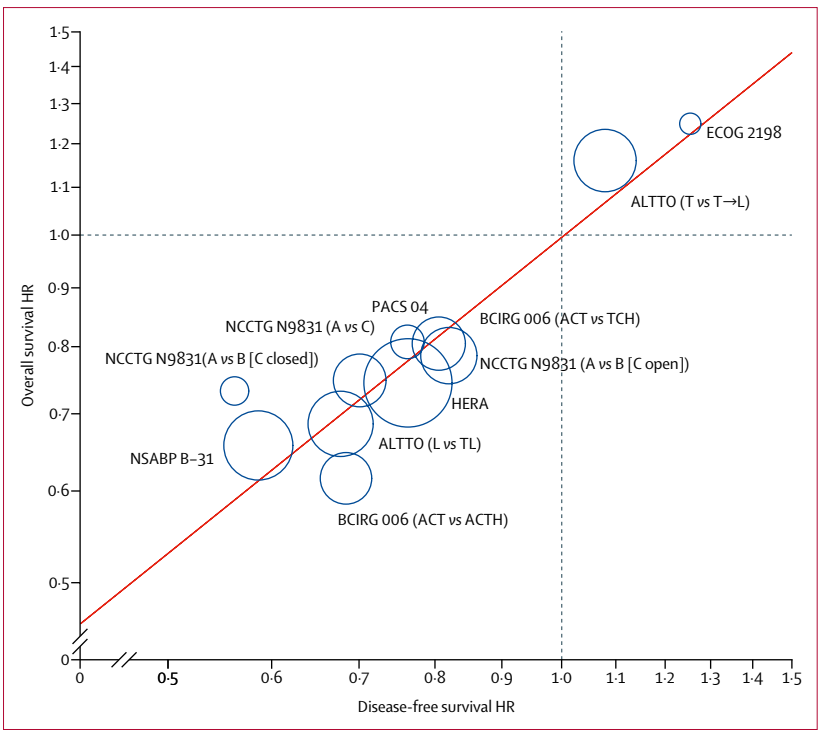

Figure 2: **Trial-level association between the hazard ratio for disease-free survival and the hazard ratio for overall survival in each contrast of the reduced set**

Analysis is weighted by the number of deaths in each contrast. Each circle represents one contrast, with size proportional to the number of deaths. ACT=doxorubicin, cyclophosphamide, and docetaxel. ACTH=doxorubicin, cyclophosphamide, docetaxel, and trastuzumab. HR=hazard ratio. L=lapatinib. T=trastuzumab. TCH=docetaxel, carboplatin, and trastuzumab. T→L=trastuzumab followed by lapatinib. TL=trastuzumab plus lapatinib.

proportional to the number of patients contributing to the contrast).

To assess model accuracy, a leave-one-out cross-validation strategy was used, with each contrast left out once and the linear model refitted to the remaining contrasts. This model was reapplied to the left-out contrast to compare the predicted and observed treatment effect on overall survival. Finally, the surrogate threshold effect was investigated. The surrogate threshold effect is the minimum treatment effect on the surrogate required to predict a non-zero treatment effect on the final endpoint in a future randomised trial.[28] Unrealistically large or small values of the surrogate threshold effect, compared with treatment effects on the surrogate observed in previous clinical trials, indicate poor validity of the surrogate.[28]

All analyses were done with SAS (version 9.4).

Role of the funding source

This study was designed and done by the authors. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication. For all trials included in the analysis, data were shared with authors from the leading institution, the International Drug Development Institute (EDS, PS, EQ, TB and MB). The costs associated with literature search and with data collection, management and analysis were defrayed by the financial support provided by Roche Pharma AG, Germany. The funder was given the opportunity to provide courtesy review of the manuscript, whose final content is the responsibility of the authors alone. The funder played no part otherwise.

Results

Of 3075 trials identified from clinical trial registries and 1320 publications identified from bibliographic databases, eight trials with available and mature data met eligibility criteria, and were included in analyses (figure 1, table 1). There was no concern about data duplication, because each trial was independent. The trial E2198[26] did not fulfil all selection criteria, because in the experimental group (as defined in this analysis) trastuzumab was planned for 12 weeks more than the usual 1 year. This was considered a minor deviation, and a decision was made by the authors to include the trial in the analysis.

Data were available from a total of 21 480 patients (3233 overall survival events and 5371 disease-free survival events) from the eight trials[2,3,4,7,9,25–27] analysed. The earliest start of accrual time for these trials was August, 1999, and the latest end of accrual time was May, 2012. The eight trials gave rise to 12 contrasts (table 1). Individual trials had slightly varying definitions for disease-free survival; because in many cases no separate information was available on non-invasive recurrences, in this analysis disease-free survival refers to any type of recurrence (invasive or non-invasive) or death from any cause. Methodological issues were identified for one of the eight eligible trials, with a much shorter follow-up and smaller number of events than all other trials.[7] The disease-free survival and overall survival curves presented in the publication of this trial had different follow-up times and distributions of censored observations between the two treatment groups, as a result of differential exclusion of randomised patients during the first year (10 in one group and 2 in the other). These methodological issues precluded proper ITT analysis, so we repeated some analyses post-hoc in a reduced set (seven trials and 11 contrasts) that excluded this study.

Table 2 shows selected results from each contrast; these results might differ from those in the original publications owing to the use of different contrasts from those used in the original trial or longer follow-up at the time of this analysis. Eight of the 12 contrasts were comparisons of chemotherapy or observation versus the same plus 12 months of trastuzumab. In three contrasts, the comparison was between 12 months of trastuzumab and shorter durations of the antibody, combined with chemotherapy alone or chemotherapy plus lapatinib. Finally, one contrast was the comparison of lapatinib for 12 months versus lapatinib plus trastuzumab for 12 months.

Patient-level associations between disease-free survival and overall survival were strong: in both the prespecified full and the post-hoc reduced sets, the r_s value was 0·90 (95% CI 0·89–0·90) in both cases. The Kaplan-Meier curves for disease-free survival and overall survival in the experimental and control groups defined for the current analyses are shown in the appendix (p 7).

For trial-level associations between disease-free survival and overall survival, in the prespecified full set, analyses weighted by the number of deaths, owing to lack of model fittting, gave rise to an R^2 value of 0·75 (95% CI 0·50–1·00), whereas in the reduced set the value of R^2 was 0·84 (95% CI 0·67–1·00). Figure 2 displays the linear regression model weighted by the number of deaths in the post-hoc reduced set, which yielded the following regression equation:

$$\ln(\text{HROS}) = -0{\cdot}005 + 0{\cdot}910 \times \ln(\text{HRDFS})$$

where OS=overall survival and DFS=disease-free survival, with standard errors of the intercept estimated as 0·042 and of the slope estimated as 0·124. The regression analysis in the prespecified full set is shown in the appendix (p 8). In neither case could the regression model be fitted with adjustment for the size of the estimation errors of the treatment effects on disease-free survival and overall survival with use of a measurement-error model. The 95% CIs for R^2 were relatively wide. Cross-validation done in both the prespecified full and post-hoc reduced sets showed that only the observed HR for overall survival for the excluded trial[7] fell outside the prediction interval (table 3). Moreover, except for the NCCTG N9831[25] A versus B (C closed) contrast, the conclusion regarding the significance of the treatment effect on overall survival resulting from the 95% CI for the predicted HR always agreed with the conclusion based on the 95% CI for the observed HR.

Secondary analyses done in subgroups defined by lymph node status and hormone-receptor status led to qualitatively similar results to those found for the analyses in the prespecified full and post-hoc reduced sets (appendix p 2). In the subset of trials that had trastuzumab for 1 year in at least one experimental group, and at least one non-anti-HER2-containing group, with the same chemotherapy in both groups, seven contrasts could be formed with a total of 11 309 patients and 2248 deaths. In this analysis, the trial-level association was weaker than in the analyses in the prespecified full and post-hoc reduced sets, with an R^2 of 0·46 (95% CI 0·00–1·00), possibly because of the exclusion of contrasts with more extreme HRs. In an analysis excluding 1082 patients with previous use of neoadjuvant therapy, which was allowed in only three trials, the trial-level associations were very similar to those in the analyses in the prespecified full and post-hoc reduced sets (data not shown).

Since the prespecified full and post-hoc reduced sets were analysed by regression models weighted by the number of deaths, the surrogate threshold effect was computed for different scenarios on the basis of the expected number of deaths in a future trial. As shown in table 4, HRs for disease-free survival below 0·81 would predict significant gains in overall survival in a randomised trial with approximately 800 deaths, whereas HRs for disease-free survival below 0·66 would

	Observed HR for overall survival (95% CI)	Predicted HR for overall survival (95% CI)	Observed HR within prediction 95% CI?
Prespecified full set (12 contrasts)			
ALTTO (lapatinib *vs* TL)[9]	0·69 (0·56–0·84)	0·71 (0·57–0·88)	Yes
ALTTO (trastuzumab *vs* T→L)[9]	1·16 (0·94–1·43)	0·94 (0·72–1·24)	Yes
BCIRG 006 (ACT *vs* ACTH)[4]	0·62 (0·48–0·79)	0·72 (0·57–0·91)	Yes
BCIRG 006 (ACT *vs* TCH)[4]	0·81 (0·63–1·03)	0·82 (0·64–1·06)	Yes
E2198[26]	1·25 (0·68–2·29)	1·21 (0·64–2·30)	Yes
HERA[2]	0·74 (0·64–0·86)	0·79 (0·68–0·93)	Yes
HORG[7]	1·45 (0·57–3·67)	0·66 (0·32–1·34)	No
NCCTG N9831 (A *vs* B [C open])[3,25]	0·79 (0·62–1·00)	0·84 (0·66–1·06)	Yes
NCCTG N9831 (A *vs* C)[3,25]	0·75 (0·58–0·96)	0·72 (0·56–0·93)	Yes
NCCTG N9831 (A *vs* B [C closed])[3,25]	0·73 (0·46–1·16)	0·58 (0·38–0·91)	Yes
NSABP B-31[3]	0·66 (0·54–0·79)	0·59 (0·47–0·75)	Yes
PACS04[27]	0·81 (0·55–1·19)	0·78 (0·53–1·15)	Yes
Reduced set (11 contrasts)			
ALTTO (L *vs* TL)[9]	0·69 (0·56–0·84)	0·70 (0·59–0·84)	Yes
ALTTO (trastuzumab *vs* T→L)[9]	1·16 (0·94–1·43)	0·95 (0·77–1·17)	Yes
BCIRG 006 (ACT *vs* ACTH)[4]	0·62 (0·48–0·79)	0·71 (0·60–0·86)	Yes
BCIRG 006 (ACT *vs* TCH)[4]	0·81 (0·63–1·03)	0·82 (0·66–1·01)	Yes
E2198[26]	1·25 (0·68–2·29)	1·22 (0·71–2·08)	Yes
HERA[2]	0·74 (0·64–0·86)	0·79 (0·69–0·90)	Yes
NCCTG N9831 (A *vs* B [C open])[3,25]	0·79 (0·62–1·00)	0·84 (0·69–1·02)	Yes
NCCTG N9831 (A *vs* C)[3,25]	0·75 (0·58–0·96)	0·72 (0·58–0·88)	Yes
NCCTG N9831 (A *vs* B [C closed])[3,25]	0·73 (0·46–1·16)	0·58 (0·41–0·81)	Yes
NSABP B-31[3]	0·66 (0·54–0·79)	0·58 (0·49–0·70)	Yes
PACS04[27]	0·81 (0·55–1·19)	0·78 (0·56–1·07)	Yes

HR=hazard ratio. ACT=doxorubicin, cyclophosphamide, and docetaxel. ACTH=doxorubicin, cyclophosphamide, docetaxel, and trastuzumab. TCH=docetaxel, carboplatin, and trastuzumab. T→L=trastuzumab followed by lapatinib. TL=trastuzumab plus lapatinib.

Table 3: **Cross-validation using leave-one-out analysis**

	Prespecified full set (12 contrasts)	Post-hoc reduced set (11 contrasts)
100	0·56	0·59
200	0·66	0·69
400	0·75	0·77
800	0·81	0·83

Numbers are HR for disease-free survival.

Table 4: **Surrogate threshold effects in the full and reduced sets, according to the expected number of deaths in a future randomised trial**

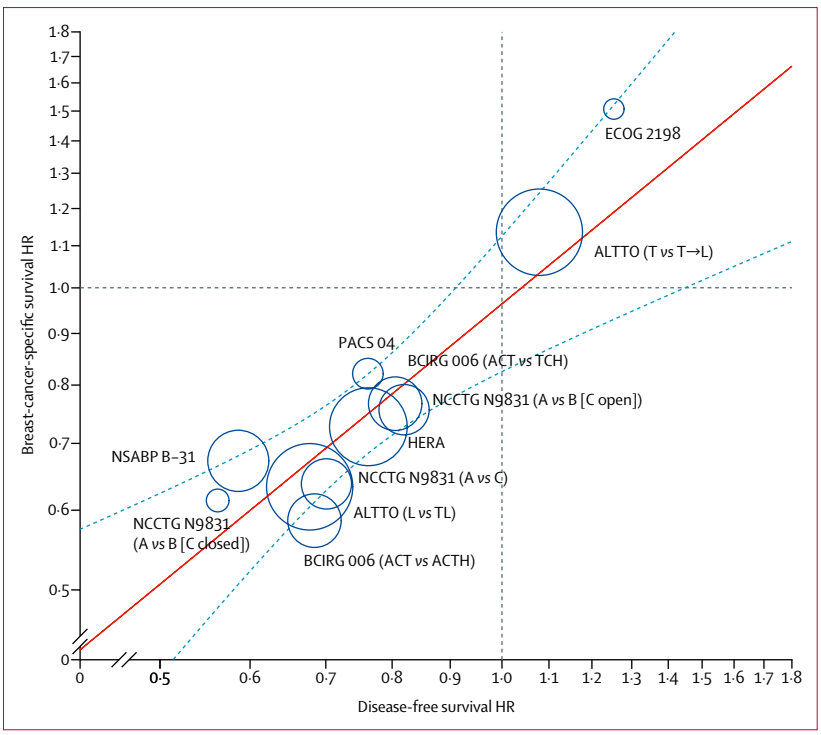

Figure 3: **Trial-level association between the hazard ratio for disease-free survival and the hazard ratio for a proxy to breast cancer-specific survival in each contrast of the reduced set**
Analysis is adjusted for the magnitude of the estimation errors in the treatment effect estimates. Each circle represents one contrast, with size proportional to the number of deaths. The curved diagonal lines are the 95% prediction limits for the regression line. ACT=doxorubicin, cyclophosphamide, and docetaxel. ACTH=doxorubicin, cyclophosphamide, docetaxel, and trastuzumab. HR=hazard ratio. L=lapatinib. T=trastuzumab. TCH=docetaxel, carboplatin, and trastuzumab. T→L=trastuzumab followed by lapatinib. TL=trastuzumab plus lapatinib.

	Observed HR for DFS	Observed HR for OS (95% CI)	Predicted HR for OS (95% CI)
PHARE[6]*	1·28 (1·05–1·56)	1·46 (1·06–2·01)	1·25 (0·93–1·67)
Short-HER[20]†	1·15 (0·91–1·46)	1·06 (0·73–1·55)	1·13 (0·78–1·64)
SOLD[21]†	1·39 (1·12–1·72)	1·36 (0·98–1·89)	1·34 (0·94–1·91)

Predicted HRs obtained using the regression equation for the main analysis. DFS=disease-free survival. OS=overall survival. *HRs with 6 months of trastuzumab as experimental group and 12 months as control group. †HRs with 9 weeks of trastuzumab as experimental group and 12 months as control group.

Table 5: **Comparison between observed and predicted hazard ratios (HRs) for overall survival based on published HR for disease-free survival in three trials with unavailable data**

predictably be followed by significant gains in overall survival in trials with approximately 200 deaths.

In prespecified exploratory analyses, we evaluated the association between disease-free survival and breast-cancer specific survival using the 2744 deaths occurring after a recurrence (85% of the total 3233 deaths). In this case, the results for the reduced set could be obtained from a model adjusted for the magnitude of the estimation errors of the treatment effects by use of a measurement-error model. As a result, the 95% prediction limits could be estimated in a uniform way (ie, irrespective of the number of deaths; figure 3). In this analysis, the patient-level association was marked by the r_s value of 0·98 (95% CI 0·98–0·98), whereas the trial-level association was characterised by an R^2 of 0·95 (95% CI 0·60–1·00) and by the following regression equation:

$$\ln(\text{HRBCSS}) = -0{\cdot}037 + 0{\cdot}929 \times \ln(\text{HRDFS})$$

where BCSS=breast cancer-specific survival and DFS=disease-free survival with standard errors of the intercept estimated as 0·073 and of the slope estimated as 0·237. Other results from this exploratory analysis are shown in the appendix (p 2). Finally, disease-free survival results from three trials with unavailable data as of this writing were used to verify some overall survival predictions based on the reduced-set model described previously (table 5). In all cases, the observed HR for overall survival is included in the prediction interval and the conclusion regarding the significance of the treatment effect on overall survival resulting from the prediction interval agrees with the conclusion based on the 95% CI for the observed HR.

Discussion

The results of this systematic review and meta-analysis suggest a strong association between disease-free survival and overall survival, and between the treatment effects on these two endpoints, when assessing adjuvant trastuzumab in patients with early HER2-positive breast cancer. The measure of treatment-level association (R^2) was equal to or above 0·75 in both the full and the post-hoc reduced sets analysed herein; this value is a commonly used threshold for accepting the validity of a surrogate for overall survival.[19,29] However, the 95% CIs for these estimates were wide, with the lower limit of the CI less than 0·75 in both analysis sets, precluding any definitive conclusions. Using Lassere's criteria for evaluation of surrogate endpoints,[29] disease-free survival had good statistical validity as a surrogate for overall survival in our study, with individual-level R^2=0·81 (0·92), trial-level $R^2 \geq 0{\cdot}75$ and a surrogate threshold effect ≥0·75 in a future trial large enough to observe about 400 deaths (table 4). Unfortunately, models taking into account the size of the estimation error in the estimated treatment effects on disease-free survival and overall survival could not be fitted for the analyses having overall survival as the final endpoint, which might have led to biased estimates of the strength of the association between the treatment effects. Conversely, in an exploratory analysis the use of the model that adjusted for the size of the estimation error in the treatment estimates was possible for breast cancer-specific survival as the final endpoint. This model yielded an R^2 value of 0·95, lending additional support to the notion that recurrences are on the causal pathway to death in early breast cancer.

Despite the biological rationale for considering disease-free survival as a surrogate for overall survival in

early breast cancer, patients with breast cancer are often older and die from other causes. Therefore, there is a need to confirm whether disease-free survival and overall survival are associated both at the patient and at the trial levels. To our knowledge, this study is the first to formally assess disease-free survival as a surrogate for overall survival in the adjuvant treatment of HER2-positive breast cancer. In a study presented in abstract form, data from nearly 12 000 patients enrolled in five phase 3 trials were used to assess various potential surrogates for overall survival in the adjuvant treatment of breast cancer.[30] Two of the trials analysed in that study are also included here,[2,27] but the others did not assess anti-HER2 therapy. The authors of that study[30] found that invasive disease-free survival had the strongest association with overall survival at the trial level but concluded that further evaluation on a larger set of trials was required to improve the precision of their estimations. Moreover, it is conceivable that the association between endpoints differs according to breast cancer phenotype, as suggested in a meta-analysis in the neoadjuvant setting.[14] For this reason, we believe that our results pertain to HER2-positive disease, and separate studies should be done for luminal and triple-negative phenotypes of breast cancer. Of note, most trials enrolling patients with these phenotypes have been designed on the basis of treatment type and not the phenotype, and (with few exceptions) patients with luminal and triple-negative disease represent subgroups among the totality of enrolled patients in trials of hormone therapy and chemotherapy. Obtaining specific data from those patients is a foreseeable difficulty in future surrogacy work related to the HER2-negative phenotypes.

The main limitation of this study is that regression analyses are very sensitive to outliers. In this study, exclusion of one trial with short follow-up and few events, and with different censoring patterns during the first year of follow-up in some of the analyses, led to quantitatively different results from those in the full set. Nevertheless, the values of R^2 equal to or above 0·75 in the full and reduced sets are reassuring in this regard, although the lower limit of the 95% CIs of these values fall below the 0·75 threshold. Additional limitations exist, one of which relates to our definition of trial eligibility. At the time that the study was planned, and still to this date in several countries, 1 year of trastuzumab was the standard of care for HER2-positive disease. As a result, our findings cannot be expanded to different settings, such as longer treatment with trastuzumab or extension of adjuvant therapy through the use of neratinib.[10] Another limitation relates to our analysis of breast cancer-specific survival, which has emerged as an endpoint more recently. Since breast cancer-specific survival had not been assessed systematically in most of the trials analysed here, we used as a proxy the cases of death preceded by a recurrence. Thus, the analysis of breast cancer-specific survival is exploratory, and it will be important to compare the predictive ability of disease-free survival and breast cancer-specific survival when both have been collected systematically in a sufficient number of trials. Although we did not consider selection bias to be an issue because all studies included in this analysis were randomised, observation bias cannot be excluded in the case of disease-free survival, because all trials were unblinded. Finally, the small number of contrasts in this study precludes meaningful analyses in subsets defined by different patient subgroups (eg, node-positive, node-negative, hormone receptor-positive, etc, and combinations of these) or trial types (eg, trastuzumab used alone, with chemotherapy, or with lapatinib). Thus, given our inability to differentiate disease-free survival from invasive disease-free survival owing to the heterogeneity of definitions across trials, our surrogacy results apply to disease-free survival broadly defined, but future studies should try to compare the predictive ability of these two endpoints. Similarly, future studies should investigate separately the trials that compared 1 year of trastuzumab versus other durations of the antibody, most of which are unavailable to us at present.[6,7,9,20–22]

Our results are in line with those obtained in colon, gastric, and non-small-cell lung cancer, in which disease-free survival was found to be an acceptable surrogate for overall survival in the adjuvant setting.[16–18] Similar conclusions were drawn for relapse-free survival in the adjuvant therapy of melanoma.[31] In those studies, the estimated values of R^2 ranged from 0·91 to 0·96 in their respective main analyses. Such values suggest stronger correlations between treatment effects in those settings than found here. Whether this is due to specific features of the trials analysed, biological differences between these clinical settings, or the play of chance, remains speculative. Arguably, the efficacy of anti-HER2 therapy in the metastatic setting, and the fact that patients with early breast cancer often die from other causes, might attenuate the association between treatment effects in breast cancer, by comparison with other settings. The analyses with a proxy for breast cancer-specific survival, which showed higher values of R^2 than the analyses with overall survival as the final endpoint, provide indirect support for this argument.

In summary, our results suggest that disease-free survival might have good overall statistical validity to be used as a surrogate for overall survival in the adjuvant treatment of HER2-positive, early breast cancer. These results, which apply mainly to the adjuvant use of trastuzumab for 12 months, indicate levels of association, both at the patient and at the trial level, that are promising from the point of view of replacing a final endpoint such as overall survival.

Contributors

All authors participated in study design through drafting or approval of the protocol. EDS contributed to the literature search. EDS, PS, EQ, SD,

DM, EP, MP-G, BPS, DS, and NW worked on data collection. PS, TB, and MB analysed the data. EDS, PS, TB, SD, DM, EP, MP-G, BPS, DS, NW, and MB interpreted the data. EDS, TB, and MB wrote the manuscript draft. All authors reviewed and approved the final version of the manuscript.

Declaration of interests

EDS and PS report grants from Roche Pharma AG to the International Drug Development Institute (their employer) during the conduct of the study. SD reports grants, personal fees, and non-financial support from Roche-Genentech, and grants and personal fees from Puma, outside of the submitted work. MP-G reports personal fees from Roche-Genentech, outside of the submitted work. DS reports consultancy for Novartis, Eli Lilly, and Seattle Genetics, and is a board member at BioMarin. NW reports grants from the National Cancer Institute, during the conduct of the study. MB is an employee and holds stock at the International Drug Development Institute. All other authors declare no competing interests.

Acknowledgments

We thank Pascale Holzerny and Johannes Ziemen, from HealthEcon (Basel, Switzerland), for doing the systematic search.

References

1 Curigliano G, Burstein HJ, Winer EP, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann Oncol* 2017; **28:** 1700–12.

2 Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 2005; **353:** 1659–72.

3 Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005; **353:** 1673–84.

4 Slamon D, Eiermann W, Robert N, et al. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med* 2011; **365:** 1273–83.

5 Joensuu H, Kellokumpu-Lehtinen PL, Bono P, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med* 2006; **354:** 809–20.

6 Pivot X, Romieu G, Debled M, et al. 6 months versus 12 months of adjuvant trastuzumab for patients with HER2-positive early breast cancer (PHARE): a randomised phase 3 trial. *Lancet Oncol* 2013; **14:** 741–48.

7 Mavroudis D, Saloustros E, Malamos N, et al. Six versus 12 months of adjuvant trastuzumab in combination with dose-dense chemotherapy for women with HER2-positive breast cancer: a multicenter randomized study by the Hellenic Oncology Research Group (HORG). *Ann Oncol* 2015; **26:** 1333–40.

8 Cameron D, Piccart-Gebhart MJ, Gelber RD, et al. 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. *Lancet* 2017; **389:** 1195–205.

9 Piccart-Gebhart M, Holmes E, Baselga J, et al. Adjuvant lapatinib and trastuzumab for early human epidermal growth factor receptor 2-positive breast cancer: results from the randomized phase III adjuvant lapatinib and/or trastuzumab treatment optimization Trial. *J Clin Oncol* 2016; **34:** 1034–42.

10 Chan A, Buyse M, Yao B. Neratinib after trastuzumab in patients with HER2-positive breast cancer—author's reply. *Lancet Oncol* 2016; **17:** e176–77.

11 von Minckwitz G, Procter M, de Azambuja E, et al. Adjuvant pertuzumab and trastuzumab in early her2-positive breast cancer. *N Engl J Med* 2017; **377:** 122–31.

12 National Comprehensive Cancer Network. NCCN clinical practice guidelines in oncology. Breast cancer, version 4. 2017. http://www.nccn.org/professionals/physician_gls/default.aspx (accessed March 16, 2018).

13 Burzykowski T, Saad ED, Buyse M. Adoption of pathologic complete response as a surrogate end point in neoadjuvant trials in HER2-positive breast cancer still an open question. *JAMA Oncol* 2017; **3:** 416.

14 Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet* 2014; **384:** 164–72.

15 Gill S, Sargent D. End points for adjuvant therapy trials: has the time come to accept disease-free survival as a surrogate end point for overall survival? *Oncologist* 2006; **11:** 624–29.

16 Sargent DJ, Wieand HS, Haller DG, et al. Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials. *J Clin Oncol* 2005; **23:** 8664–70.

17 Oba K, Paoletti X, Alberts S, et al. Disease-free survival as a surrogate for overall survival in adjuvant trials of gastric cancer: a meta-analysis. *J Natl Cancer Inst* 2013; **105:** 1600–07.

18 Mauguen A, Pignon JP, Burdett S, et al. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. *Lancet Oncol* 2013; **14:** 619–26.

19 Burzykowski T, Molenberghs G, Buyse M. The evaluation of surrogate endpoints. New York: Springer, 2005.

20 Conte PF, Bisagni G, Frassoldati A, et al. 9 weeks *vs* 1 year adjuvant trastuzumab in combination with chemotherapy: results of the phase III multicentric Italian study Short-HER. *J Clin Oncol* 2017; **35** (suppl)**:** 501–01.

21 Joensuu H, Fraser J, Wildiers H, et al. A randomized phase III study of adjuvant trastuzumab for a duration of 9 weeks versus 1 year, combined with adjuvant taxane-anthracycline chemotherapy, for early HER2-positive breast cancer (the SOLD study). *Cancer Res* 2018; **78** (suppl)**:** GS3-04.

22 Earl HM, Hiller L, Vallier A-L, et al. PERSEPHONE: 6 versus 12 months (m) of adjuvant trastuzumab in patients (pts) with HER2 positive (+) early breast cancer (EBC): randomised phase 3 non-inferiority trial with definitive 4-year (yr) disease-free survival (DFS) results. *J Clin Oncol* 2018; **36** (suppl 15)**:** 506 (abstr).

23 Singhal MK, Kapoor A, Narayan S, et al. A phase III randomized study of paclitaxel and trastuzumab versus paclitaxel alone for early stage, ER and PR receptor negative and HER2-positive breast cancer as adjuvant treatment. *Ann Oncol* 2015; **26** (suppl)**:** iii1–iii3.

24 Gourgou-Bourgade S, Cameron D, Poortmans P, et al. Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials). *Ann Oncol* 2015; **26:** 873–79.

25 Perez EA, Suman VJ, Davidson NE, et al. Sequential versus concurrent trastuzumab in adjuvant chemotherapy for breast cancer. *J Clin Oncol* 2011; **29:** 4491–97.

26 Schneider BP, O'Neill A, Shen F, et al. Pilot trial of paclitaxel-trastuzumab adjuvant therapy for early stage breast cancer: a trial of the ECOG-ACRIN cancer research group (E2198). *Br J Cancer* 2015; **113:** 1651–57.

27 Spielmann M, Roche H, Delozier T, et al. Trastuzumab for patients with axillary-node-positive breast cancer: results of the FNCLCC-PACS 04 trial. *J Clin Oncol* 2009; **27:** 6129–34.

28 Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat* 2006; **5:** 173–86.

29 Lassere MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res* 2008; **17:** 303–40.

30 Savina M, Jacot W, Mathoulin-Pélissier S, et al. Surrogate endpoints for overall survival in randomized controlled trials evaluating adjuvant treatment for breast cancer: a meta-analysis. *Ann Oncol* 2017; **28** (suppl 5)**:** v43–v67.

31 Suciu S, Eggermont AMM, Lorigan P, et al. Relapse-free survival as a surrogate for overall survival in the evaluation of stage ii-iii melanoma adjuvant therapy. *J Natl Cancer Inst* 2018; **110:** 87–96.