

Intermediate clinical endpoints in early-stage breast cancer: an analysis of individual patient data from the Gruppo Italiano Mammella and Mammella Intergruppo trials



Eva Blondeaux,^{a,*} Wanling Xie,^b Luca Carmisciano,^c Silvia Mura,^d Valeria Sanna,^d Michelino De Laurentiis,^e Roberta Caputo,^e Anna Turletti,^f Antonio Durando,^g Sabino De Placido,^h Carmine De Angelis,^h Giancarlo Bisagni,ⁱ Elisa Gasparini,ⁱ Anita Rimanti,^j Fabio Puglisi,^{k,l} Mauro Mansutti,^m Elisabetta Landucci,ⁿ Alessandra Fabi,^o Luca Arecco,^{p,q} Marta Perachino,^{p,q} Marco Bruzzone,^q Luca Boni,^q Matteo Lambertini,^{p,q} Lucia Del Mastro,^{p,q,r} and Meredith M. Regan^{b,r}



^aU.O. Epidemiology Unit, IRCCS Ospedale Policlinico San Martino, Genoa, Italy

^bDivision of Biostatistics, Dana-Farber Cancer Institute, Boston, USA

^cDipartimento di Medicina Clinica e Sperimentale, Università di Pisa, Pisa, Italy

^dDepartment of Medical Oncology, UOC Oncologia Medica, University Hospital of Sassari, Sassari, Italy

^eIstituto Nazionale per lo Studio e la Cura dei Tumori, Fondazione Pascale IRCCS, Napoli, Italy

^fMedical Oncology, Ospedale Martini ASL Città di Torino, Torino, Italy

^gBreast Unit, Città della Salute e della Scienza, Ospedale S. Anna, Torino, Italy

^hDepartment of Clinical Medicine and Surgery, University of Naples Federico II, Napoli, Italy

ⁱDepartment of Oncology and Advanced Technology, Oncology Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy

^jASST Mantova, Azienda Ospedaliera Carlo Poma, Mantova, Italy

^kDepartment of Medical Oncology, Unit of Medical Oncology and Cancer Prevention, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, Aviano, Italy

^lDepartment of Medicine (DAME), University of Udine, Udine, Italy

^mAzienda Sanitaria Universitaria Friuli Centrale, Udine, Italy

ⁿMedical Oncology, S. Chiara Hospital, Pisa, Italy

^oPrecision Medicine Breast Unit, Scientific Directorate, Department of Women, Children and Public Health Sciences, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

^pDepartment of Medical Oncology, U.O. Clinica di Oncologia Medica, IRCCS Ospedale Policlinico San Martino, Genoa, Italy

^qDepartment of Internal Medicine and Medical Sciences (DiMI), School of Medicine, University of Genova, Genova, Italy

Summary

Background Intermediate clinical endpoints (ICEs) are frequently used as primary endpoint in randomised trials (RCTs). We aim to assess whether changes in different ICEs can be used to predict changes in overall survival (OS) in adjuvant breast cancer trials.

Methods Individual patient level data from adjuvant phase III RCTs conducted by the Gruppo Italiano Mammella (GIM) and Mammella Intergruppo (MIG) study groups were used. ICEs were computed according to STEEP criteria. Using a two-stage meta-analytic model, we assessed the surrogacy of each ICE at both the outcome (i.e., OS and ICE are correlated irrespective of treatment) and trial (i.e., treatment effects on ICE and treatment effect on OS are correlated) levels. The following ICEs were considered as potential surrogate endpoints of OS: disease-free survival (DFS), distant disease-free survival (DDFS), distant relapse-free survival (DRFS), recurrence-free survival (RFS), recurrence-free interval (RFI), distant recurrence-free interval (DRFI), breast cancer-free interval (BCFI), and invasive breast cancer-free survival (IBCFS). The estimates of the degree of correlation were obtained by copula models and weighted linear regression. Kendall's τ and $R^2 \geq 0.70$ were considered as indicators of a clinically relevant surrogacy.

Findings Among the 12,397 patients enrolled from November 1992 to July 2012 in six RCTs, median age at enrolment was 57 years (interquartile range (IQR) 49–65). After a median follow-up of 10.3 years (IQR 6.4–14.5), 2131 (17.2%) OS events were observed, with 1390 (65.2%) attributed to breast cancer. At the outcome-level, Kendall's τ ranged from 0.69 for BCFI to 0.84 for DRFS. For DFS, DDFS, DRFS, RFS, RFI, DRFI, BCFI, and IBCFS endpoints, over 95% of the 8-year OS variability was attributable to the variation of the 5-year ICE. At the trial-level, treatment effects for the different ICEs and OS were strongly correlated, with the highest correlation for RFS and DRFS and the lowest for BCFI.

eClinicalMedicine

2024;70: 102501

Published Online 20 March 2024

<https://doi.org/10.1016/j.eclinm.2024.102501>

*Corresponding author. U.O. Epidemiology Unit, IRCCS Ospedale Policlinico San Martino, Largo Rosanna Benzi 10, Genoa 16132, Italy.

E-mail address: eva.blondeaux@hsanmartino.it (E. Blondeaux).

^rContributed equally.

Interpretation Our results provide evidence supporting the use of DFS, DDFS, DRFS, RFS, RFI, DRFI, and IBCFS as primary endpoint in breast cancer adjuvant trials.

Funding This analysis was supported by the Italian Association for Cancer Research (“Associazione Italiana per la Ricerca sul Cancro”, AIRC; IG 2017/20760) and by Italian Ministry of Health–5 × 1000 funds (years 2021–2022).

Copyright © 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Intermediate clinical endpoint; Overall survival; Surrogate endpoint; Adjuvant setting

Research in context

Evidence before this study

The aim of the standardized definitions for efficacy endpoints (STEEP) criteria is to reduce the inconsistencies among definition of intermediate clinical endpoints in adjuvant breast cancer trials.

We searched Medline with no language or date restriction on November 15th, 2023 by using the search terms “early stage breast cancer”, “surrogate endpoint”, and “overall survival” to search for studies evaluating the surrogacy of different intermediate clinical endpoints for overall survival. Few studies assessed surrogacy of disease-free survival for overall survival: one previous meta-analysis failed to demonstrate a correlation between 2-year disease-free survival and 5-year overall survival, while a previous meta-analysis established disease-free survival as surrogate endpoint for overall in trials of adjuvant trastuzumab for patients with HER2-positive breast cancer. No studies assessing whether changes in invasive breast cancer-free survival and in other intermediate clinical endpoints can be used to predict changes in overall survival were found.

Therefore, considered the very limited evidence, assessing surrogacy of different intermediate clinical endpoints for overall survival in adjuvant breast cancer trials is needed.

Added value of this study

Over the past years, the Gruppo Italiano Mammella (GIM) and Mammella Intergruppo (MIG) study groups conducted several randomized trials in the adjuvant breast cancer setting. By pooling individual patient data from six randomized trials (n = 12,397 patients) with a median follow-up exceeding 10 years and by using the surrogate endpoint methodology to

assess correlation between different intermediate clinical endpoints and overall survival, we showed that disease-free survival, distant disease-free survival, distant relapse-free survival, recurrence-free survival, recurrence-free interval, distant recurrence-free interval, and invasive breast cancer-free survival are able to predict changes in overall survival. In the subgroup analysis restricted to patients with hormone-receptor positive/HER2-negative tumours, none of the intermediate clinical endpoints met the criteria to be considered a surrogate endpoint for overall survival. To our knowledge, this is the only study that assess surrogacy of different intermediate clinical endpoints for overall survival in adjuvant breast cancer trials.

Implications of all the available evidence

Because demonstrating an improvement in overall survival in randomized trial requires the inclusion of a substantial number of patients and long-term follow-up data, using intermediate clinical endpoint as primary endpoint in randomized trial is highly attractive with reduced number of patients and length of follow up. However, changes in intermediate clinical endpoints should be able to predict changes in overall survival. Our analysis provides evidence supporting the use of disease-free survival, distant disease-free survival, distant relapse-free survival, recurrence-free survival, recurrence-free interval, distant recurrence-free interval, and invasive breast cancer-free survival, defined by STEEP criteria v2.0, as primary endpoint in breast cancer adjuvant trials. In conclusion, these results have potential implication in trial design and drug approval.

Introduction

Improving the overall survival (OS) of patients with cancer should be considered the main goal of anticancer treatments. In randomised trials (RCTs), OS definition (i.e., the elapsed time from randomisation to death) is unique and OS is the preferred endpoint for regulatory purposes.^{1–3} However, showing OS improvements in RCTs usually require the inclusion of a substantial number of patients and long-term follow-up data. In the early stage breast cancer setting, intermediate clinical endpoints (ICEs), such as disease-free survival (DFS) or

invasive-DFS (iDFS), are frequently used as primary endpoint in RCTs and OS is often included as secondary endpoint.^{4–8}

Before the introduction of the standardized definitions for efficacy endpoints (STEEP) criteria, the definition of DFS and other ICEs was lacking consistency among different RCTs.⁹ Thus, the use of the STEEP criteria was proposed with the aim to reduce the inconsistencies that hamper the interpretation of the results across trials. In STEEP v1.0 criteria, iDFS was recognized as a potential surrogate endpoint for OS and

its use was considered justifiable due to the extended life expectancy of breast cancer patients, including patients experiencing metastatic recurrence following adjuvant treatments.⁹ In the STEEP criteria v2.0, nine ICEs were identified: DFS, iDFS, distant disease-free survival (DDFS), distant relapse-free survival (DRFS), recurrence-free survival (RFS), recurrence-free interval (RFI), distant recurrence-free interval (DRFI), breast cancer-free interval (BCFI) and invasive breast cancer-free survival (IBCFS).¹⁰ IBCFS definition, added in STEEP v2.0, exclude second primary non-breast cancers compared to iDFS definition.¹⁰ According to simulations, IBCFS might be the preferred endpoint in some circumstances such as when the intervention is expected not to increase secondary non-breast cancers.

In adjuvant breast cancer setting, there were prior attempts of establishing ICEs, especially DFS, as surrogate of OS. One previous meta-analysis failed to demonstrate, at a trial-level, a correlation between 2-year DFS and 5-year OS.¹¹ A systematic review from the Early Breast Cancer Trialists' Collaborative Group showed that improvement in DFS might predict improvements in OS.¹² However, no formal statistical analysis were undertaken to directly establish this correlation or to assess the degree of correlation.¹² To date, no attempts were made to understand how IBCFS, and other ICEs, improvements might predict OS benefits.

Over the past years, the Gruppo Italiano Mammella (GIM) and Mammella Intergruppo (MIG) study groups conducted several RCTs in the adjuvant breast cancer setting. By pooling individual patient data, we aimed to assess whether changes in IBCFS and in other ICEs can be used to predict changes in OS in adjuvant breast cancer RCTs.

Methods

Description of included studies

This analysis included individual patient-level data from six adjuvant phase III RCTs by the MIG and GIM study groups (MIG1,¹³ MIG5,¹⁴ GIM2,¹⁵ GIM3,¹⁶ GIM4,¹⁷ GIM6¹⁸). Details are described in the [Supplementary methods](#) section ([Appendix](#) pp. 2) and summarized in [Supplementary Table S1](#) ([Appendix](#) pp. 4). Briefly, all trials included were multicentre studies investigating different adjuvant chemotherapy or endocrine treatments. Results of the treatment effect according to different ICEs in MIG and GIM trials, except for ICE defined in each protocol, were not previously published. The Independent Review Boards of all participating centres approved each trial and all included patients provided written informed consent before study entry. The MIG and GIM Steering Committees approved the present analysis before its conduction. Trials enrolment spanned from November 1992 to July 2012.

Events definition

ICEs were computed according to the STEEP criteria definition.¹⁰ Briefly, the endpoints included different combinations of local-regional, distant, contralateral, second non-breast malignancies, and deaths as events ([Appendix](#) pp. 5). Notably, five ICEs (i.e., DFS, DDFS, RFS, DRFS, IBCFS) included death from any cause as survival event while three ICEs (i.e., RFI, DRFI, BCFI) included only death from breast cancer as survival event. Thus, RFS differs from RFI and DRFS differs from DRFI only for non-breast cancer related deaths. In case no survival event was recorded, patients were censored at the last follow-up visit.

When the type of a new primary breast cancer (invasive vs *in situ*) was not available in the original electronic case report form (eCRF) (i.e., from GIM3, GIM4 and GIM6 trials), it was assumed to be invasive. Because of the assumption that all new primary breast cancer in the latter trials were invasive, DFS was chosen as ICE because in this endpoint both recurrences (i.e., invasive and *in situ*) would be included as events. Thus, iDFS, that differs from DFS just by not including *in situ* recurrences as events, was not tested as a surrogate endpoint of OS.

Statistical analysis

Given the clinical relevance of lymph node status in treatment guidelines and clinical trial eligibility determination on the basis of lymph node status, trials were split into units according to nodal status.¹⁹ To avoid exclusion of patients with missing information on nodal stage, single imputation, assuming monotone missing patterns and using the logistic regression method, was performed among 75 (0.6%) patients. Variables used for imputation of nodal status were trial, arm, age and receipt of (neo)adjuvant chemotherapy.

We evaluated the surrogacy of the different ICE and OS using a meta-analytic two-stage validation model.^{19,20} Two conditions must be satisfied to claim for OS surrogacy: the outcome-level surrogacy (the ICE and OS are correlated irrespective of treatment) and the trial-level surrogacy (the treatment effects on both endpoints are correlated). We defined a priori a clinically relevant surrogacy of an R^2 value of ≥ 0.7 .^{21,22}

The outcome-level surrogacy was tested at both the patient level and trial level. At the patient level, associations of OS with different ICE were evaluated via a bivariate model fitted on individual patient data. Clayton, Hougaard and Plackett's copula functions were considered to account for individual dependence of the outcomes; copula selection was based on model goodness of fitness evaluated through the Akaike Information Criterion. Kendall's τ (range 0–1) quantified the correlation between ICE and OS at the patient level. We defined an a priori threshold of clinically relevant surrogacy at Kendall's τ value of ≥ 0.7 .^{21,22} Further details regarding copula models are provided in the

[Supplementary methods](#) section ([Appendix](#) pp. 3). For this analysis, individual data from patients enrolled in the GIM2 trial were considered once. At the trial level, we used the Kaplan–Meier method to obtain 5-year ICE estimate and 8-year OS estimate within each unit. We chose these time points as they are both frequently reported in the literature and used for trial assumption in sample size estimations. We performed weighted linear regression (WLR) analyses between units-specific OS rates at 8-year vs 5-year ICE rates. Regressions were weighted by inverse variances of the 5-year estimates of the ICE. R^2 was used to quantify the proportion of the 8-year OS rate variance that was explained by each 5-year ICE rate. Due to the 2×2 factorial design, for this analysis, individual data from patients enrolled in the GIM2 trial were considered twice: the first comparison between epirubicin and cyclophosphamide regimen (standard) vs fluorouracil, epirubicin and cyclophosphamide regimen (experimental) and the second comparison between standard interval schedule (standard) vs dose dense schedule (experimental). For this latter comparison, 88 patients enrolled in five centres providing only standard interval chemotherapy schedule were excluded.^{15,23}

For the trial-level surrogacy, we obtained the nodal- and study-specific treatment effects using Cox models. We then performed a WLR between the estimate of treatment effect on OS log (hazard ratio [HR]_{OS}) and the estimates of treatment effect on each ICE (logHR_{ICE}). Models were weighted by inverse variances of the logHR_{ICE}, and R^2 was used to quantify the proportion of variance that was explained by the regressor. For this analysis, individual data from patients enrolled in the GIM2 trial were considered twice.

Given the inclusion of only patients with hormone-receptor positive tumours in two trials, we expected most of the patients to have hormone-receptor positive/HER2-negative tumours. Thus, we planned a subgroup analysis to evaluate if changes in different ICEs can be used to predict changes in OS in the hormone-receptor positive/HER2-negative population only.

Surrogate threshold effect and validation

Surrogate threshold effect (STE) is defined as the minimum treatment effect on the surrogate endpoint necessary to predict a nonzero treatment effect—that is an HR different from 1—on OS in a future trial.²⁴ To estimate STE we constructed the 95% prediction limits for the regression line of the effect of treatment on OS vs the effect of treatment on the surrogate. STE was defined as the point of the intersection of the upper 95% prediction limit with the horizontal line representing an HR of 1 for OS.

For trial-level surrogacy, model accuracy was assessed by leave-one-out cross validation. Each unit was left out once and the WLR, performed on the remaining $n-1$ units, was used to estimate the treatment effect (log

[HR]) on OS based on the observed log (HR) of the ICE of the left out unit. R^2 was also calculated from the remaining $n-1$ units model to evaluate the impact of a single unit on the correlation between treatment effects on ICE and OS.

All analyses were performed using SAS Version 9.4 (SAS Institute, Cary, NC, USA) and R version 4.3.0 (2023-04-21; R Foundation for Statistical Computing, Vienna, Austria).

Role of the funding source

The funders of the study had no role in study design, data collection, analysis, interpretation, or writing of the report, and they had no access to the data. EB, LC, LB and LDM had full access to the data and had final responsibility for the decision to submit for publication.

Results

This analysis included 12,397 patients enrolled in six RCTs between November 1992 and July 2012. Baseline characteristics are reported in the [Supplementary materials](#) ([Appendix](#) pp. 6–7). Median age at enrolment was 57 years (Interquartile range (IQR) 49–65). Overall, 8209 (66.2%) of the patients had node positive disease, 7718 (62.3%) had hormone-receptor positive/HER2-negative tumours. Median follow-up was 10.3 years (IQR 6.4–14.5; [Appendix](#) pp. 8).

A total of 2131 (17.2%) OS events were observed, with 1390 (65.2%) attributed to breast cancer ([Appendix](#) pp. 9). Estimated survival and hazard function by years since random assignment for each ICE are shown in [Fig. 1A](#) and [B](#).

The trials were split according to nodal status and treatment arm into 22 units ([Appendix](#) pp. 10), which corresponds to 11 units for treatment comparisons.

Outcome-level surrogacy: correlation between ICE and OS irrespective of treatment

At the individual patient level, the correlation of each ICE with OS, measured through Kendall's τ , ranged from 0.69 (95% CI 0.68–0.71) for BCFI to 0.84 (95% CI 0.84–0.85) for DRFS with a Kendall's τ of 0.77 for IBCFS ([Table 1](#)).

At the trial level, a tight correlation between the nodal-, arm- and trial-specific 8-year OS estimate vs all the different ICEs was observed ($R^2 \geq 0.95$ for all ICEs) ([Table 1](#) and [Fig. 2A–H](#)).

Trial-level surrogacy: correlation between treatment effect on ICE and OS

Forest plot of observed pair of HR of treatment effect on ICE and OS endpoints are reported in [Supplementary Fig. S1A–H](#) ([Appendix](#) pp. 14–17). At the trial-level, a correlation between treatment effect on OS and on the surrogate endpoint was observed for all ICEs, with the strongest association for RFS and DRFS ($R^2 = 0.88$) and

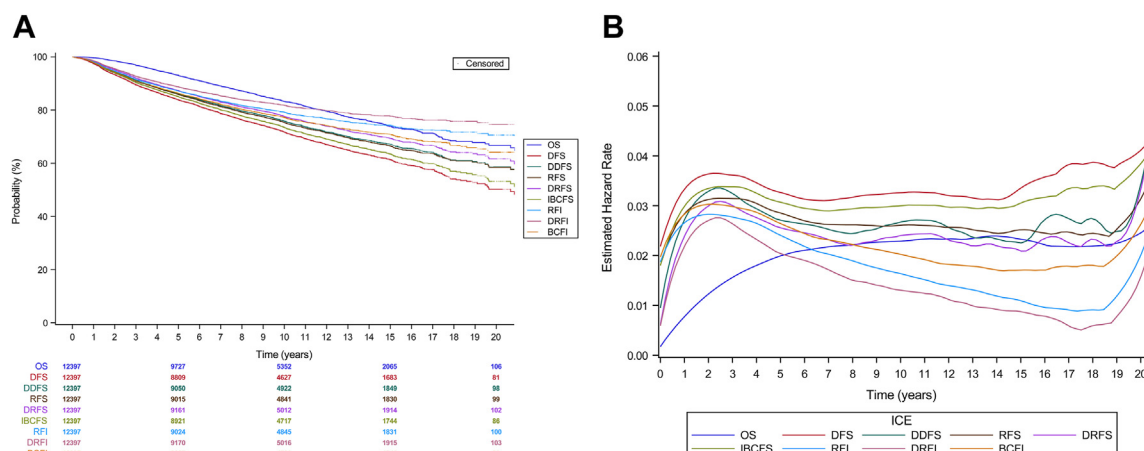


Fig. 1: (A) Kaplan-Meier estimates of OS and intermediate clinical endpoints (ICEs). (B) Estimated hazard across times for OS and intermediate clinical endpoints (ICEs). OS, overall survival; DFS, disease-free survival; DDFS, distant disease-free survival; RFS, recurrence-free survival; DRFS, distant relapse-free survival; IBCFS, invasive breast cancer-free survival; RFI, recurrence-free interval; DRFI, distant recurrence-free interval; BCFI, breast cancer-free interval.

the weakest for BCFI ($R^2 = 0.70$) (Table 1 and Fig. 3A–H). Treatment effect on IBCFS endpoint was correlated to treatment effect on OS ($R^2 = 0.84$). Similarly to patient level surrogacy, the correlation was higher for ICEs that included death from any cause as event (i.e., $R^2 \geq 0.82$), while the correlation was weaker (i.e., R^2 ranging from 0.70 to 0.77) for ICEs that excluded death from causes other than breast.

Surrogate threshold effect and validation

STE for ICEs that included death from any cause ranged from 0.85 to 0.86 while STE for ICE that included only death from breast cancer as event was approximately around 0.76 (Appendix pp. 11). Thus, a larger treatment

effect on these latter endpoints would be required to predict a treatment benefit on OS.

Results of the trial-level leave one out cross validation provided consistent results (Appendix pp. 18–22).

Hormone-receptor positive/HER2-negative population

Nodal-, arm- and trial-specific units among the hormone-receptor positive/HER2-negative population are reported in Supplementary Table S8 (Appendix pp. 12). Among the 7718 patients with known hormone-receptor positive/HER2-negative disease, median age at randomisation was 60 years (IQR 52–67; Appendix pp. 13). Overall, 1015 (13.2%) OS events were

ICE	Outcome-level surrogacy (OS and ICE are correlated irrespective of treatment)				Trial-level surrogacy (treatment effects on both end points are correlated)	
	Correlation at the patient level		Regression of 8-year OS rate v 5-year ICE rate by trial, arm, and nodal status (No. of units = 22)		Regression of Log (HR)-OS v Log (HR)-ICE by trial and nodal status (No. of units = 11)	
	No. of events out of 10,394 patients included	Kendall's τ (95% CI)	No. of events out of 12,397 patients included	R^2 (95% CI)	R^2 (95% CI)	Regression equation
DFS	2773	0.75 (0.73–0.76)	3526	0.95 (0.89–0.97)	0.82 (0.42–0.90)	$\text{Log (HR)}_{OS} = -0.056 + 1.179 * \text{Log (HR)}_{DFS}$
DDFS	2346	0.82 (0.81–0.82)	3007	0.95 (0.88–0.97)	0.86 (0.51–0.92)	$\text{Log (HR)}_{OS} = -0.036 + 1.141 * \text{Log (HR)}_{DDFS}$
RFS	2392	0.80 (0.79–0.81)	3053	0.96 (0.92–0.98)	0.88 (0.59–0.93)	$\text{Log (HR)}_{OS} = -0.023 + 1.124 * \text{Log (HR)}_{RFS}$
DRFS	2163	0.84 (0.84–0.85)	2779	0.96 (0.92–0.98)	0.88 (0.59–0.93)	$\text{Log (HR)}_{OS} = -0.023 + 1.037 * \text{Log (HR)}_{DRFS}$
IBCFS	2595	0.77 (0.76–0.78)	3306	0.97 (0.92–0.98)	0.84 (0.47–0.91)	$\text{Log (HR)}_{OS} = -0.046 + 1.117 * \text{Log (HR)}_{IBCFS}$
RFI	1853	0.73 (0.72–0.74)	2437	0.96 (0.91–0.97)	0.76 (0.31–0.87)	$\text{Log (HR)}_{OS} = -0.018 + 0.942 * \text{Log (HR)}_{RFI}$
DRFI	1582	0.77 (0.76–0.79)	2117	0.95 (0.89–0.97)	0.77 (0.31–0.87)	$\text{Log (HR)}_{OS} = -0.014 + 0.846 * \text{Log (HR)}_{DRFI}$
BCFI	2073	0.69 (0.68–0.71)	2709	0.96 (0.91–0.97)	0.70 (0.20–0.83)	$\text{Log (HR)}_{OS} = -0.048 + 0.958 * \text{Log (HR)}_{BCFI}$

DFS, disease-free survival; HR, hazard ratio; ICE, intermediate clinical end point; OS, overall survival; DDFS distant disease-free survival; RFS, recurrence-free survival; DRFS, distant relapse-free survival; RFI, recurrence-free interval; DRFI, distant recurrence-free interval; BCFI, breast cancer-free interval; IBCFS, invasive breast cancer-free survival.

Table 1: Two-condition surrogacy analysis among all patients.

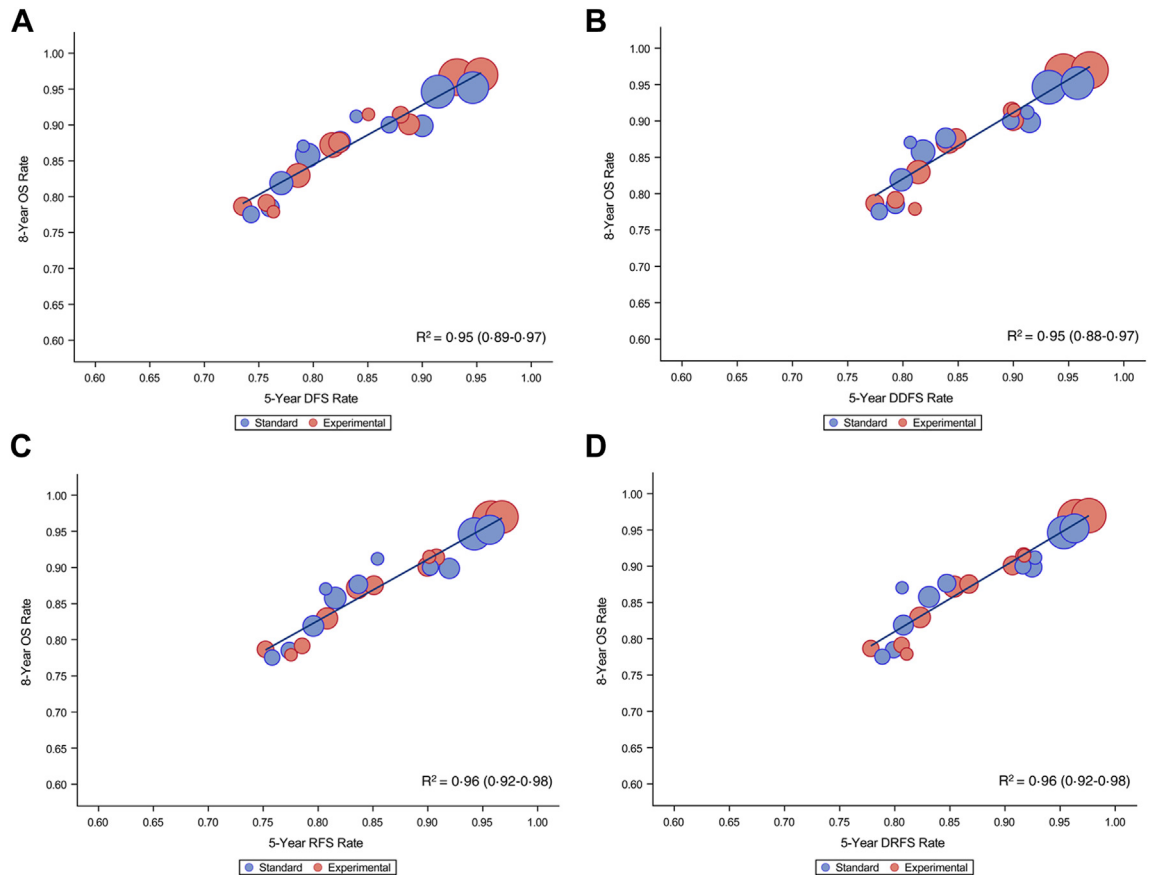


Fig. 2: Overall survival (OS) rate at 8-year vs intermediate clinical endpoints rate at 5 years: (A) 8-year OS vs 5-year DFS; (B) 8-year OS vs 5-year DDFS; (C) 8-year OS vs 5-year RFS; (D) 8-year OS vs 5-year DRFS; (E) 8-year OS vs 5-year IBCFS; (F) 8-year OS vs 5-year RFI; (G) 8-year OS vs 5-year DRFI; (H) 8-year OS vs 5-year BCFI. OS, overall survival; DFS, disease-free survival; DDFS, distant disease-free survival; RFS, recurrence-free survival; DRFS, distant relapse-free survival; IBCFS, invasive breast cancer-free survival; RFI, recurrence-free interval; DRFI, distant recurrence-free interval; BCFI, breast cancer-free interval.

observed, with 616 (60.7%) attributed to breast cancer (Appendix pp. 9).

At the outcome-level surrogacy, at the individual patient level, the correlation of each ICE with OS, measured through Kendall’s τ , was higher than 0.70 for all ICEs, except for BCFI (Kendall’s $\tau = 0.67$), with IBCFS demonstrating a strong correlation (Kendall’s $\tau = 0.77$) (Table 2). A strong correlation between the nodal-, arm- and trial-specific 8-year OS estimate vs all the different ICEs was observed (Table 2; Appendix pp. 23–26). However, at the trial-level surrogacy, a weaker correlation between log (HR)-OS and log (HR)-ICE across nodal- and trial units was observed for all ICEs (R^2 for IBCFS = 0.49) (Table 2; Appendix pp. 27–30).

Discussion

By using individual patient data from six RCTs with a median follow-up exceeding 10 years, we showed that DFS, DDFS, DRFS, RFS, RFI, DRFI, and IBCFS, are

able to predict changes in OS. ICEs that included death from any cause as event presented the strongest correlation with OS while the correlation was weaker for ICEs that censored patients with death from causes other than breast. For patients with hormone receptor-positive/HER2-negative breast cancer, none of the ICEs proposed met the criteria to be considered a surrogate endpoint of OS.

To reduce the number of patients to include and the lengths of the follow-up, DFS or iDFS have been used as primary endpoint in the majority of phase III RCTs in the breast cancer adjuvant setting.^{4–6,25} Understanding performance and utility of different ICEs is highly relevant as ICEs are frequently used in trial design and IBCFS could be considered as an alternative to iDFS in some situations. In this framework we used surrogate endpoint methodology to assess correlation between different ICEs and OS. We demonstrated that, at outcome-level, the correlation at the patient level was demonstrated for all ICEs except for BCFI (Kendall’s

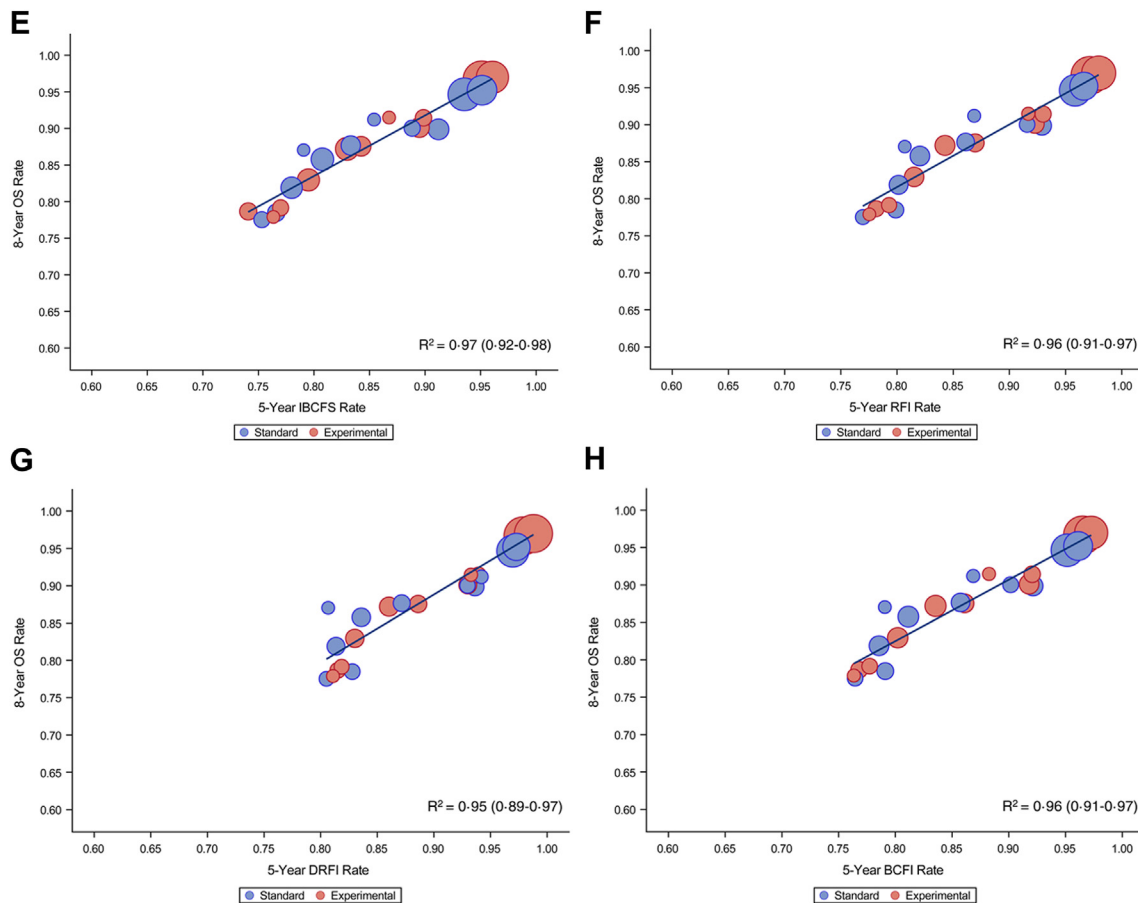


Fig. 2: (continued).

$\tau = 0.69$), with endpoints including death from any cause presenting the strongest correlation, while the association of 5-year ICE vs 8-year OS was strong (i.e., $R^2 \geq 0.95$) for all ICE. The effect of treatment on the ICEs and on the true endpoint (i.e., in the trial-level surrogacy) was equal to or above the threshold of 0.70 needed to claim for surrogacy for all ICEs tested. From a drug development point of view, the trial level association is the most interesting as it allows to estimate the treatment effect on OS given the observed treatment effect on the ICE. However, caution is needed as in some situation outcome level surrogacy is fulfilled, while attempt to demonstrate trial level surrogacy fail.²⁶ Furthermore, the use of ICEs in trial design allows an earlier observation of the events, increased power for detection of treatment effect, reduced number of patients to be included in the trial and consequently reduced costs. However, 95% CIs for the R^2 estimates were wide, with the lower limit of the 95% CI less than 0.70 for all ICEs, precluding any definitive conclusions. Among the different ICEs, DRFS presented the strongest correlation with OS. This could be explained by the

fact that distant relapses are often incurable and precede death while other events, such as contralateral breast cancer or second non-breast malignancies, could be treated with curative intent. In fact, the stronger correlation of OS with ICE that include death from any cause as a survival event could be explained by greater similarity in survival events between the two endpoints. Furthermore, IBCFS, the new ICE proposed in STEEP criteria v2.0, presented a strong correlation with OS.

In a setting of constant improvement in breast cancer care, patients with early-stage disease might die from other causes. Indeed, in a population with a median age at enrolment of 57 years and 10-year median follow-up, we found that nearly 35% events were not related to breast cancer or were due to an unknown cause. RFS and DRFS presented a stronger correlation with OS both at the outcome-level (Kendall's $\tau = 0.80$ and 0.84 for RFS and DRFS respectively) and at the trial-level ($R^2 = 0.88$ for both ICEs). The correlation was weaker for RFI and DRFI (patient level Kendall's $\tau = 0.73$ and 0.77 and trial level $R^2 = 0.76$ and 0.77 for RFI and DRFI, respectively). The lower values of Kendall's τ and R^2 coefficients

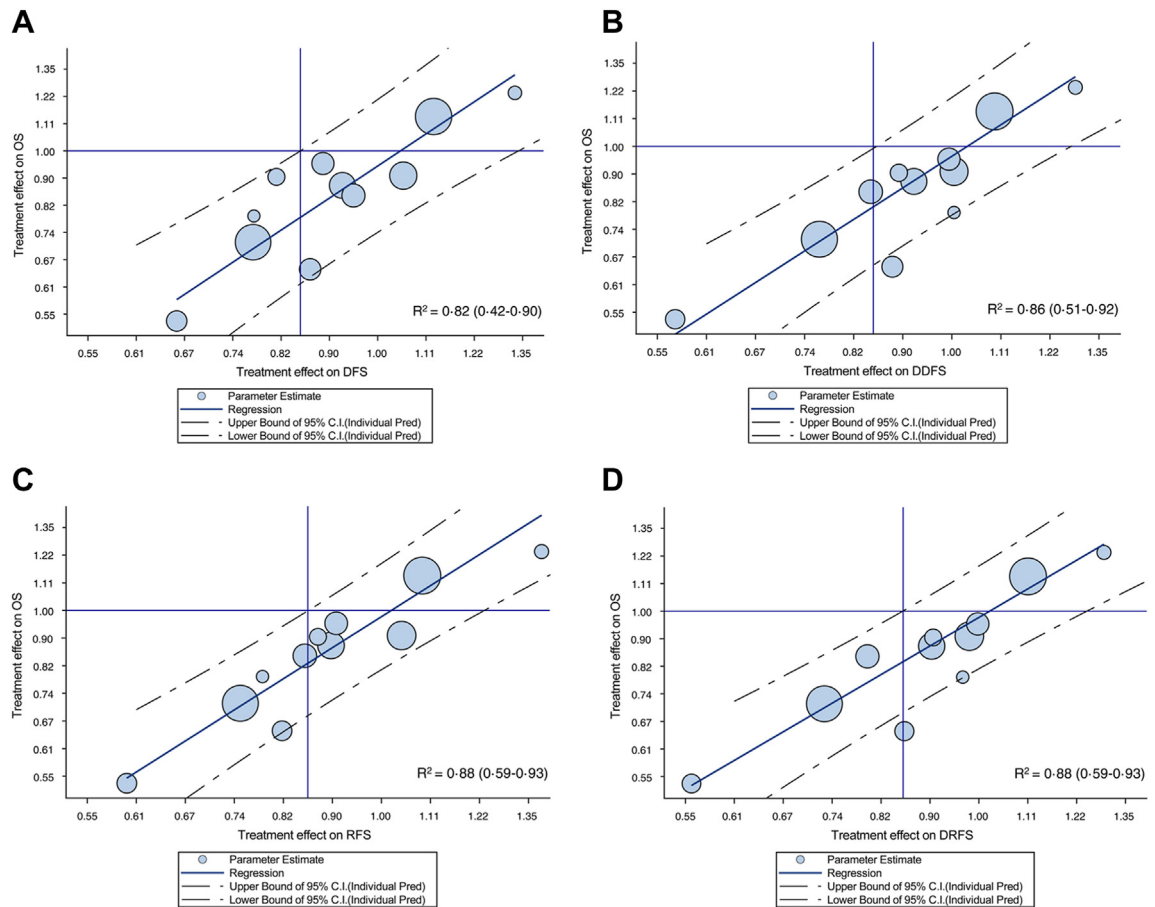


Fig. 3: Treatment effects (log hazard ratio [HR]) on overall survival (OS) vs treatment effects on intermediate clinical endpoints. (A) Log (HR)OS vs log (HR)DFS; (B) log (HR)OS vs log (HR)DDFS; (C) log (HR)OS vs log (HR)RFS; (D) log (HR)OS vs log (HR)DRFS; (E) log (HR)OS vs log (HR)IBCFS; (F) log (HR)OS vs log (HR)RFI; (G) log (HR)OS vs log (HR)DRFI; (H) log (HR)OS vs log (HR)BCFI. OS, overall survival; DFS, disease-free survival; DDFS, distant disease-free survival; RFS, recurrence-free survival; DRFS, distant relapse-free survival; IBCFS, invasive breast cancer-free survival; RFI, recurrence-free interval; DRFI, distant recurrence-free interval; BCFI, breast cancer-free interval.

comparing RFS with RFI and DRFS with DRFI could be explained by competing cause of death. Similarly in an analysis of surrogate endpoint in prostate cancer trials, a weaker correlation between endpoints that did not include death from non-cancer causes was observed compared to the correlation observed for endpoints that included death from all causes as event.²⁷

The use of ICEs offers several advantages in clinical trial research such as the early observation of the necessary events. However, some uncertainty remains regarding the ability of ICEs to accurately predict the impact on OS. STE implies that a trial with an observed upper limit of the CI for the estimated HR on the ICE that falls below the STE would predict a significant effect on OS. The lower correlation of ICEs that censored patients with death from causes other than breast ($R^2_{(RFI)} = 0.76$, $R^2_{(DRFI)} = 0.77$) resulted in lower STE (0.77 vs 0.75 for RFI and DRFI respectively) compared to ICEs that included death from all cause as events

(STE ranging from 0.85 to 0.86). Thus, a greater treatment effect is needed to be observed on ICEs that censored patients with death from causes other than breast to predict an OS benefit.

Given the potential heterogeneity of the included population in the trials and the suggestion from a previous meta-analysis that the association between endpoints could differ according to breast cancer subtypes,²⁸ we performed an analysis restricted to patients with known hormone-receptor positive/HER2-negative tumours. Despite observing a tight correlation in the outcome-level surrogacy, we failed to demonstrate trial-level surrogacy for all ICEs. This is consistent with previous trials of adjuvant endocrine treatment where, despite a strong treatment effect on ICE, little or no treatment effect on OS was demonstrated.^{4,5,29} Other possible explanations include that a lower number of patients was included and consequently a lower number of events observed. Moreover, patients with hormone-

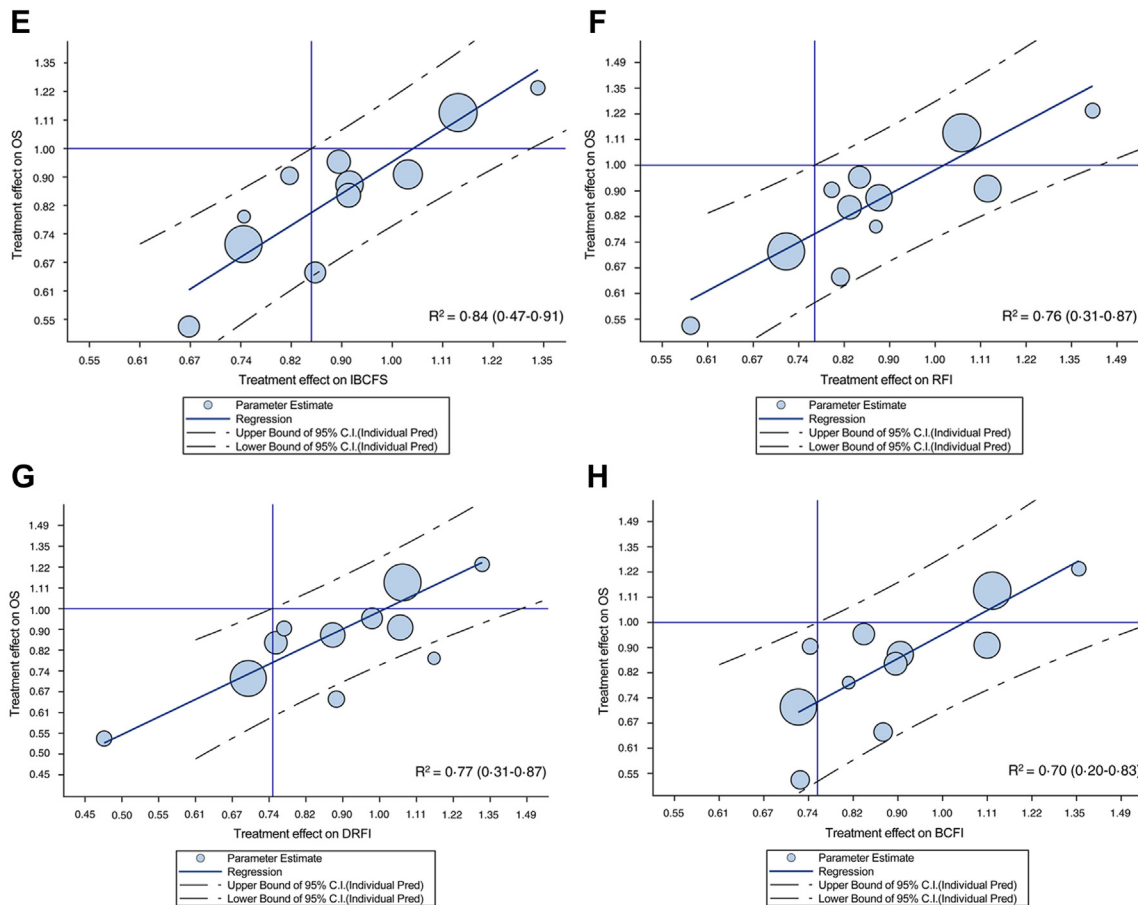


Fig. 3: (continued).

ICE	Outcome-level surrogacy (OS and ICE are correlated irrespective of treatment)				Trial-level surrogacy (treatment effects on both end points are correlated)	
	Correlation at the patient level		Regression of 8-year OS rate v 5-year ICE rate by trial, Arm, and nodal status (No. of units = 16)		Regression of Log (HR)-OS v Log (HR)-ICE by trial and nodal status (No. of units = 8)	
	No. of events out of 6612 patients included	Kendall's τ (95% CI)	No. of events out of 7718 patients included	R ² (95% CI)	R ² (95% CI)	Regression equation
DFS	1450	0.73 (0.72-0.75)	1855	0.92 (0.79-0.95)	0.54 (0.00-0.76)	$\text{Log (HR)}_{\text{OS}} = -0.158 + 0.666 * \text{Log (HR)}_{\text{DFS}}$
DDFS	1211	0.80 (0.79-0.82)	1566	0.92 (0.79-0.95)	0.67 (0.05-0.83)	$\text{Log (HR)}_{\text{OS}} = -0.110 + 0.811 * \text{Log (HR)}_{\text{DDFS}}$
RFS	1204	0.80 (0.79-0.82)	1558	0.94 (0.83-0.96)	0.66 (0.04-0.82)	$\text{Log (HR)}_{\text{OS}} = -0.109 + 0.756 * \text{Log (HR)}_{\text{RFS}}$
DRFS	1082	0.85 (0.84-0.86)	1413	0.94 (0.84-0.96)	0.67 (0.05-0.83)	$\text{Log (HR)}_{\text{OS}} = -0.082 + 0.791 * \text{Log (HR)}_{\text{DRFS}}$
IBCFS	1325	0.77 (0.75-0.78)	1707	0.94 (0.84-0.96)	0.49 (0.00-0.74)	$\text{Log (HR)}_{\text{OS}} = -0.147 + 0.617 * \text{Log (HR)}_{\text{IBCFS}}$
RFI	900	0.71 (0.69-0.73)	1206	0.95 (0.87-0.97)	0.49 (0.00-0.74)	$\text{Log (HR)}_{\text{OS}} = -0.114 + 0.531 * \text{Log (HR)}_{\text{RFI}}$
DRFI	767	0.76 (0.74-0.78)	1048	0.94 (0.85-0.97)	0.53 (0.00-0.75)	$\text{Log (HR)}_{\text{OS}} = -0.090 + 0.539 * \text{Log (HR)}_{\text{DRFI}}$
BCFI	1030	0.67 (0.65-0.69)	1365	0.95 (0.85-0.97)	0.29 (0.00-0.63)	$\text{Log (HR)}_{\text{OS}} = -0.158 + 0.400 * \text{Log (HR)}_{\text{BCFI}}$

DFS, disease-free survival; HR, hazard ratio; ICE, intermediate clinical end point; OS, overall survival; DDFS distant disease-free survival; RFS, recurrence-free survival; DRFS, distant relapse-free survival; RFI, recurrence-free interval; DRFI, distant recurrence-free interval; BCFI, breast cancer-free interval; IBCFS, invasive breast cancer-free survival.

Table 2: Two-condition surrogacy analysis among patients with hormone-receptor positive/HER2-negative breast cancer.

receptor positive/HER2-negative tumours have different baseline characteristics and different pattern of recurrence as compared to other tumour subtypes.³⁰

Due to the low number of patients known to have HER2-positive breast cancer (n = 1480, 11.9%) and the available evidence on the use of DFS as surrogate endpoint of OS in trials of adjuvant trastuzumab,³¹ the analysis in this subgroup was not performed. Similarly, only a minority of patients had triple-negative breast cancer; hence, no subgroup testing could be planned.

Several limitations of our study should be acknowledged. Firstly, we used data only from adjuvant trials conducted by the GIM and MIG collaborative groups and no attempt to perform a systematic review or to obtain individual data from other collaborative group was made. The included trials spanned the accrual period from 1992 to 2012; considering the recent major advances in the adjuvant breast cancer treatment, the applicability of these results to modern RCTs may be limited. Included trials comprised different therapeutic strategies (chemotherapy and endocrine therapy) and the small number of nodal-, arm- and trial-units precluded subgroup analyses by nodal status or trial type. On the other hand, the clear results using heterogeneous trials is a strength of the generalizability of the results. Ductal carcinoma *in situ* is a rare event and we believe that assuming for three trials that all new primary breast cancer were invasive did not bias our results. However, due to this assumption, we were not able to test iDFS endpoint. Finally, we did not account for different centre-specific effect. Nevertheless, in Italy healthcare system is based on universal coverage and provides comprehensive services to its citizens. All patients enrolled were treated in high volumes hospitals and no major differences in survival outcome could be expected.

Despite these limitations, this is the first attempt to demonstrate performances of different ICEs on predicting changes in OS in adjuvant breast cancer RCTs.

In conclusion, our study provides evidence supporting the use of DFS, DDFS, DRFS, RFS, RFI, DRFI, and IBCFS, defined by STEEP criteria v2.0, as primary endpoint in breast cancer adjuvant trials. ICEs that included death for any cause (i.e., DFS, DDFS, RFS, DRFS, IBCFS) presented strong correlation with OS. For patients with hormone receptor-positive/HER2-negative disease, none of the ICEs met the criteria to be considered a surrogate endpoint of OS. Future research should further investigate the application of ICEs in different breast cancer subtypes and in the setting of new targeted treatment strategies.

Contributors

Eva Blondeaux MD (Conceptualization; Data curation; Formal Analysis; Writing—original draft), Wanling Xie MSc (Methodology; Writing—review & editing), Luca Carmisciano MD (Formal Analysis; Writing—review & editing), Silvia Mura MD (Investigation; Writing—review &

editing), Valeria Sanna MD (Investigation; Writing—review & editing), Michelino De Laurentiis MD (Investigation; Writing—review & editing), Roberta Caputo MD (Investigation; Writing—review & editing), Anna Turletti MD (Investigation; Writing—review & editing), Antonio Durando MD (Investigation; Writing—review & editing), Sabino De Placido MD (Investigation; Writing—review & editing), Carmine De Angelis MD (Investigation; Writing—review & editing), Giancarlo Bisagni MD (Investigation; Writing—review & editing), Elisa Gasparini MD (Investigation; Writing—review & editing), Anita Rimanti MD (Investigation; Writing—review & editing), Fabio Puglisi MD (Investigation; Writing—review & editing), Mauro Mansutti MD (Investigation; Writing—review & editing), Elisabetta Landucci MD (Investigation; Writing—review & editing), Alessandra Fabi MD (Investigation; Writing—review & editing), Luca Arecco MD (Investigation; Writing—review & editing), Marta Perachino MD (Investigation; Writing—review & editing), Marco Bruzzone MSc (Formal Analysis; Methodology; Writing—review & editing), Luca Boni MD (Formal Analysis; Methodology; Writing—review & editing), Matteo Lambertini PhD (Conceptualization; Investigation; Supervision; Writing—review & editing), Lucia Del Mastro MD (Conceptualization; Investigation; Writing—review & editing), Meredith M Regan ScD (Conceptualization; Formal Analysis; Methodology; Supervision; Writing—original draft).

Data sharing statement

Individual participant data that underlies the results reported in this article will be available for further sharing upon reasonable request to the corresponding author and after approval from the MIG and GIM study groups. Data will be available after publication and ending 5 years following article publication.

Declaration of interests

Eva Blondeaux reports research support (to the Institution) from Gilead outside the submitted work.

Wanling Xie served as consultant to Convergent Therapeutics, Inc outside the submitted work.

Carmine De Angelis reports advisory role for Roche, Lilly, Novartis, AstraZeneca, Pfizer, Seagen, Daiichi-Sankyo, Gilead, and GSK and speaker honoraria from Roche, Lilly, Novartis, Pfizer, Seagen, GSK, GILEAD, and Daiichi-Sankyo. Travel Grants from Gilead and research support (to the Institution) from Novartis, GILEAD, and Daiichi-Sankyo outside the submitted work.

Fabio Puglisi reports honoraria for advisory boards, activities as a speaker, travel grants, research grants from Amgen—Astrazeneca—Daiichi Sankyo—Celgene—Eisai—Eli Lilly—Exact Sciences—Gilead—Ipsen—Menarini—MSD—Novartis—Pierre Fabre—Pfizer—Roche—Seagen—Takeda—Viartis and research funding from Astrazeneca—Eisai—Roche outside the submitted work.

Mauro Mansutti reports honoraria for advisory boards, activities as a speaker, travel grants: Accord Healthcare, Amgen, AstraZeneca, Eli Lilly, Gilead, MSD, Novartis, Pfizer, Seagen outside the submitted work.

Alessandra Fabi reports advisory board from Roche, Novartis, Lilly, Pfizer, MSD, Pierre Fabre, Eisai, Epionpharma, Gilead, Seagen, AstraZeneca, Exact Science and consultant from Dompè Farmaceutica outside the submitted work.

Matteo Lambertini reports advisory role for Roche, Lilly, Novartis, AstraZeneca, Pfizer, Seagen, Gilead, MSD and Exact Sciences and speaker honoraria from Roche, Lilly, Novartis, Pfizer, Sandoz, Libbs, Daiichi Sankyo, Knight and Takeda, Travel Grants from Gilead and Daiichi Sankyo, and research support (to the Institution) from Gilead outside the submitted work.

Lucia Del Mastro reports grants or contracts from Eli Lilly, Novartis, Roche, Daiichi Sankyo, and Seagen; honoraria from Roche, Novartis, Pfizer, Eli Lilly, AstraZeneca, MSD, Seagen, Gilead, Pierre Fabre, Eisai, Exact Sciences, and Ipsen; support for attending meetings or travel from Roche, Pfizer, and Eisai; participation on a Data Safety Monitoring Board or Advisory Board from Novartis, Roche, Eli Lilly, Pfizer, Daiichi Sankyo, Exact Sciences, Gilead, Pierre Fabre, Eisai, and AstraZeneca outside the submitted work.

Meredith M Regan reports consulting or advisory role for Ipsen (to the Institution), Tolmar, Bristol Myers Squibb, Debiopharm Group (to the Institution), TerSera, AstraZeneca and research funding from Pfizer (to the Institution), Ipsen (to the Institution), Novartis (to the Institution), Merck (to the Institution), AstraZeneca (to the Institution), Pierre Fabre (to the Institution), Bayer (to the Institution), Bristol Myers Squibb (to the Institution), Roche (to the Institution), TerSera (to the Institution), Debiopharm Group (to the Institution), BioTheragnostics (to the Institution) and honoraria from Bristol Myers Squibb and Canadian Urological Association.

Acknowledgements

The authors thank patients, physicians, nurses, and trial coordinators who participated in the trials. We acknowledge the Italian Association for Cancer Research ("Associazione Italiana per la Ricerca sul Cancro", AIRC; IG 2017/20760) and the Italian Ministry of Health (5 × 1000 funds, years 2021-2022) for the funding.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.eclim.2024.102501>.

References

- 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;11:624–629.
- Research C for DE and. E17 general principles for planning and design of multi-regional clinical trials; 2020. published online April 29. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e17-general-principles-planning-and-design-multi-regional-clinical-trials>. Accessed August 31, 2023
- EMA. Evaluation of anticancer medicinal products in man—scientific guideline. European Medicines Agency; 2018. published online Sept 17. <https://www.ema.europa.eu/en/evaluation-anticancer-medicinal-products-man-scientific-guideline>. Accessed August 31, 2023
- Mamounas EP, Bandos H, Lembersky BC, et al. A randomized trial of five years of letrozole versus placebo after aromatase inhibitor-based therapy: NRG Oncology/NSABP B-42. *Lancet Oncol*. 2019;20:88–99.
- Pagani O, Walley BA, Fleming GF, et al. Adjuvant exemestane with ovarian suppression in premenopausal breast cancer: long-term follow-up of the combined TEXT and SOFT trials. *J Clin Oncol*. 2023;41:1376–1382.
- Perrone F, De Laurentis M, De Placido S, et al. Adjuvant zoledronic acid and letrozole plus ovarian function suppression in premenopausal breast cancer: HOBEO phase 3 randomised trial. *Eur J Cancer*. 2019;118:178–186.
- Johnston SRD, Toi M, O'Shaughnessy J, et al. Abemaciclib plus endocrine therapy for hormone receptor-positive, HER2-negative, node-positive, high-risk early breast cancer (monarchE): results from a preplanned interim analysis of a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2023;24:77–90.
- Loibl S, Marmé F, Martin M, et al. Palbociclib for residual high-risk invasive HR-positive and HER2-negative early breast cancer—the penelope-B trial. *J Clin Oncol*. 2021;39:1518–1530.
- Hudis CA, Barlow WE, Costantino JP, et al. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J Clin Oncol*. 2007;25:2127–2132.
- Tolaney SM, Garrett-Mayer E, White J, et al. Updated standardized definitions for efficacy end points (STEEP) in adjuvant breast cancer clinical trials: STEEP version 2.0. *J Clin Oncol*. 2021;39:2720–2731.
- Ng R, Pond GR, Tang PA, MacIntosh PW, Siu LL, Chen EX. Correlation of changes between 2-year disease-free survival and 5-year overall survival in adjuvant breast cancer trials from 1966 to 2006. *Ann Oncol*. 2008;19:481–486.
- Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. 2005;365:1687–1717.
- Blondeaux E, Lambertini M, Michelotti A, et al. Dose-dense adjuvant chemotherapy in early breast cancer patients: 15-year results of the Phase 3 Mammella InterGruppo (MIG)-1 study. *Br J Cancer*. 2020;122:1611–1617.
- Del Mastro L, Levaggi A, Michelotti A, et al. 5-Fluorouracil, epirubicin and cyclophosphamide versus epirubicin and paclitaxel in node-positive early breast cancer: a phase-III randomized GONO-MIG5 trial. *Breast Cancer Res Treat*. 2016;155:117–126.
- Del Mastro L, Poggio F, Blondeaux E, et al. Fluorouracil and dose-dense adjuvant chemotherapy in patients with early-stage breast cancer (GIM2): end-of-study results from a randomised, phase 3 trial. *Lancet Oncol*. 2022;23:1571–1582.
- De Placido S, Gallo C, De Laurentis M, et al. Adjuvant anastrozole versus exemestane versus letrozole, upfront or after 2 years of tamoxifen, in endocrine-sensitive breast cancer (FATA-GIM3): a randomised, phase 3 trial. *Lancet Oncol*. 2018;19:474–485.
- Del Mastro L, Mansutti M, Bisagni G, et al. Extended therapy with letrozole as adjuvant treatment of postmenopausal patients with early-stage breast cancer: a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol*. 2021;22:1458–1467.
- Lambertini M, Boni L, Michelotti A, et al. Long-term outcomes with pharmacological ovarian suppression during chemotherapy in premenopausal early breast cancer patients. *J Natl Cancer Inst*. 2022;114:400–408.
- Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics*. 2000;1:49–67.
- Ciani O, Davis S, Tappenden P, et al. Validation of surrogate endpoints in advanced solid tumors: systematic review of statistical methods, results, and implications for policy makers. *Int J Technol Assess Health Care*. 2014;30:312–324.
- ICECaP Working Group. The development of intermediate clinical endpoints in cancer of the prostate (ICECaP). *J Natl Cancer Inst*. 2015;107:djv261.
- Xie W, Halabi S, Tierney JF, et al. A systematic review and recommendation for reporting of surrogate endpoint evaluation using meta-analyses. *JNCI Cancer Spectr*. 2019;3:pkz002.
- Del Mastro L, De Placido S, Bruzzi P, et al. Fluorouracil and dose-dense chemotherapy in adjuvant treatment of patients with early-stage breast cancer: an open-label, 2 × 2 factorial, randomised phase 3 trial. *Lancet*. 2015;385:1863–1872.
- Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat*. 2006;5:173–186.
- Martín M, Ruiz Simón A, Ruiz Borrego M, et al. Epirubicin plus cyclophosphamide followed by docetaxel versus epirubicin plus docetaxel followed by capecitabine as adjuvant therapy for node-positive early breast cancer: results from the GEICAM/2003-10 study. *J Clin Oncol*. 2015;33:3788–3795.
- Buyse M, Saad ED, Burzykowski T, Regan MM, Sweeney CS. Surrogacy beyond prognosis: the importance of 'trial-level' surrogacy. *Oncologist*. 2022;27:266–271.
- Xie W, Regan MM, Buyse M, et al. Metastasis-free survival is a strong surrogate of overall survival in localized prostate cancer. *J Clin Oncol*. 2017;35:3097–3104.
- 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;384:164–172.
- Goss PE, Ingle JN, Pritchard KI, et al. Extending aromatase-inhibitor adjuvant therapy to 10 years. *N Engl J Med*. 2016;375:209–219.
- Metzger-Filho O, Sun Z, Viale G, et al. Patterns of recurrence and outcome according to breast cancer subtypes in lymph node-negative disease: results from international breast cancer study group trials VIII and IX. *J Clin Oncol*. 2013;31:3083–3090.
- Saad ED, Squifflet P, Burzykowski T, et al. 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. *Lancet Oncol*. 2019;20:361–370.