Title

Datopotamab—deruxtecan plus durvalumab in early-stage breast cancer: the sequential multiple assignment randomized I-SPY2.2 phase 2 trial

Authors

Shatsky, Rebecca A
Trivedi, Meghna S
Yau, Christina
et al.

Publication Date

2024-09-14

DOI

10.1038/s41591-024-03267-1

nature medicine

Article

https://doi.org/10.1038/s41591-024-03266-2

Datopotamab–deruxtecan in early-stage breast cancer: the sequential multiple assignment randomized I-SPY2.2 phase 2 trial

Received: 7 August 2024

Accepted: 23 August 2024

Published online: 14 September 2024

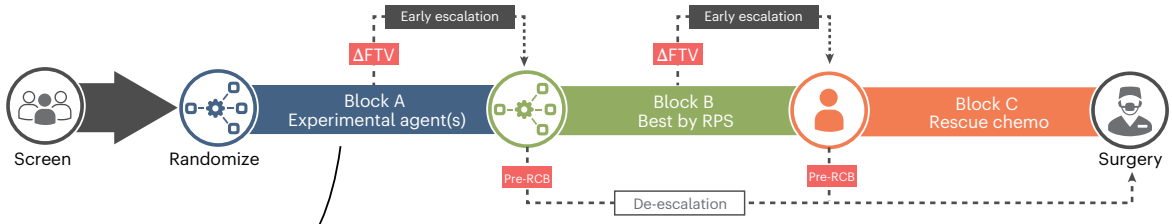
Check for updates

A list of authors and their affiliations appears at the end of the paper

Among the goals of patient-centric care are the advancement of effective personalized treatment, while minimizing toxicity. The phase 2 I-SPY2.2 trial uses a neoadjuvant sequential therapy approach in breast cancer to further these goals, testing promising new agents while optimizing individual outcomes. Here we tested datopotamab–deruxtecan (Dato-DXd) in the I-SPY2.2 trial for patients with high-risk stage 2/3 breast cancer. I-SPY2.2 uses a sequential multiple assignment randomization trial design that includes three sequential blocks of biologically targeted neoadjuvant treatment: the experimental agent(s) (block A), a taxane-based regimen tailored to the tumor subtype (block B) and doxorubicin–cyclophosphamide (block C). Patients are randomized into arms consisting of different investigational block A treatments. Algorithms based on magnetic resonance imaging and core biopsy guide treatment redirection after each block, including the option of early surgical resection in patients predicted to have a high likelihood of pathological complete response, the primary endpoint. There are two primary efficacy analyses: after block A and across all blocks for the six prespecified breast cancer subtypes (defined by clinical hormone receptor/human epidermal growth factor receptor 2 (HER2) status and/or the response-predictive subtypes). We report results of 103 patients treated with Dato-DXd. While Dato-DXd did not meet the prespecified threshold for success (graduation) after block A in any subtype, the treatment strategy across all blocks graduated in the hormone receptor-negative HER2$^-$Immune$^-$DNA repair deficiency$^-$ subtype with an estimated pathological complete response rate of 41%. No new toxicities were observed, with stomatitis and ocular events occurring at low grades. Dato-DXd was particularly active in the hormone receptor-negative/HER2$^-$Immune$^-$DNA repair deficiency$^-$ signature, warranting further investigation, and was safe in other subtypes in patients who followed the treatment strategy. ClinicalTrials.gov registration: NCT01042379.

e-mail: laura.esserman@ucsf.edu

https://doi.org/10.1038/s41591-024-03266-2



Treatment assignments/randomization

$HR^+HER2^-Immune^-DRD^-$	Taxol	AC
$HR^-HER2^-Immune^-DRD^-$	Taxol + carbo + pembro	AC + pembro
$HER2^-Immune^+$	Taxol + carbo + pembro	AC + pembro
$HER2^-Immune^-DRD^+$	Taxol + carbo + pembro	AC + pembro
$HER2^-Immune^-DRD^+$	Taxol + carbo	AC + pembro

Fig. 1 | I-SPY2.2 study schematic. Participants receive up to three rounds (or blocks) of treatment, depending upon response assessed by MRI and biopsy. Block A is a platform trial design that randomizes participants to one of several arms testing experimental agents or combinations. Patients not responding to treatment in block A receive paclitaxel-based standard of care therapy in block B, assigned (or randomized) based on their RPS. Those not responding in block B proceed to treatment in block C and receive a minimum of AC chemotherapy. All patients complete treatment with definitive surgery. Patients who meet the threshold to have a high probability of RCB 2/3 disease at 6 weeks (based on MRI at 0, 3 and 6 weeks in A or after 6 weeks in B) have the option of forgoing the remainder of treatment in that block (early treatment redirection to next block). Surgeon icon by Delwar Hossain released under a CC BY 3.0 license. Taxol, paclitaxel; carbo, carboplatin; pembro, pembrolizumab.

The I-SPY2 trial, which began enrolling patients in 2010 at sites across the United States, was designed to rapidly assess the value of adding novel drugs to the standard of care neoadjuvant chemotherapy backbone for patients with high-risk, early stage breast cancer. Patients enrolled in I-SPY2 received a 12 week course of paclitaxel (and trastuzumab ± pertuzumab, if $HER2^+$) along with one of several novel investigational agents, followed by four cycles of traditional doxorubicin–cyclophosphamide before surgery. Over 12 years, 23 drugs and drug combinations were evaluated, with some—such as pembrolizumab—going on to phase 3 trials and changing the standard of care worldwide for women with some of the highest-risk subtypes of breast cancer[1,2].

I-SPY2 enrolled over 2,100 patients, and careful study of these individuals led to several key developments that have shaped the treatment of high-risk breast cancer as well as future trial design. The trial demonstrated that pathologic complete response (pCR) is an important prognostic marker for individuals with molecularly high-risk breast cancer, regardless of the treatment they receive[3]. By evaluating patients with serial magnetic resonance imaging (MRIs) and biopsies during neoadjuvant therapy, I-SPY2 investigators developed an algorithm to predict pCR that could potentially identify a group of patients that could safely go to surgery before completing an entire planned neoadjuvant chemotherapy regimen[4]. Finally, investigators were able to classify early stage breast cancers beyond just estrogen receptor, progesterone receptor and HER2 status into response-predictive subtypes (RPS) that reflect potential sensitivity to modern classes of treatments. The RPS is defined by a combination of gene expression signatures (response to immunotherapy and/or DNA repair deficiency (DRD)) and the BluePrint assay, which could allow for further personalization of therapy in the future[5].

Using the lessons from I-SPY2, I-SPY2.2 was designed with the goal of further refining the biological targeting of treatments to give each woman the best opportunity to reach pCR, while at the same time reducing exposure to unnecessary treatments to minimize toxicities. It retains similar drug development goals as I-SPY2, in that I-SPY2.2 is a phase 2 signal-finding trial to efficiently identify targeted experimental treatments that are likely to be successful in phase 3 trials. Patients enrolled in I-SPY2.2 receive up to three sequential blocks of neoadjuvant treatment (Fig. 1). Investigational agents/combinations are evaluated in block A, which has a randomized platform design permitting the investigation of multiple agents in parallel; block B assigns patients to taxane-based chemotherapy, with additional agents (carboplatin, trastuzumab, pertuzumab and/or pembrolizumab) tailored to a specific RPS based upon standards of care and I-SPY2 experience; and block C is traditional doxorubicin–cyclophosphamide with pembrolizumab for the immune-positive and triple-negative subsets. Patients are monitored closely with serial MRI throughout each block and biopsy at the end of a block. These form the basis of an algorithm (pre-residual cancer burden (pre-RCB)) to assess the likelihood that a patient has achieved pCR and provide appropriate recommendations for treatment redirection. If pre-RCB indicates a high probability of pCR, patients are recommended to proceed to surgery (de-escalation), foregoing treatment in subsequent blocks; when pre-RCB is below a preset threshold, patients are recommended to proceed early to the next block of treatment (early escalation). This data-directed treatment redirection allows patients to avoid toxicity and overtreatment with therapies they may not need. Additionally, because in I-SPY2, patients who did not reach a specific threshold for change in MRI functional tumor volume after 6 weeks were unlikely to have a meaningful response to that regimen[6,7], I-SPY2.2 allows patients who did not meet this threshold of response on a particular block after 6 weeks to switch to the next block of treatment early. This allows patients to avoid toxicity from treatments that are unlikely to add benefit.

Here, we report the safety and efficacy results of the evaluation of datopotamab–deruxtecan (Dato-DXd), an arm in block A of I-SPY2.2.

Dato-DXd is an antibody drug conjugate (ADC) comprising a humanized antitrophoblast cell-surface antigen 2 (TROP2) IgG1 monoclonal antibody attached to a topoisomerase I inhibitor payload via a plasma-stable, cleavable linker[8]. In preclinical models, Dato-DXd has been shown to induce DNA damage and apoptosis in TROP2-expressing tumor cells. TROP2 is often highly expressed in triple-negative breast cancer (TNBC), but can be expressed in tumor cells of any breast cancer subtype, and as such has emerged as an interesting therapeutic target in metastatic breast cancer[9].

In the phase 1 TROPION-PanTumor01 study that included patients with metastatic non-small cell lung cancer, hormone receptor $(HR)^+$/human epidermal growth factor receptor 2 $(HER2)^-$ breast cancer and TNBC, Dato-DXd showed promising clinical activity with an overall response rate of 31.8% in heavily pretreated patients with TNBC, 26.8% in $HR^+/HER2^-$ breast cancer and 40% in patients with TNBC who had not received prior treatment with a topoisomerase inhibitor-based ADC[10].

A pooled analysis of Dato-DXd for metastatic TNBC demonstrated a favorable safety profile, with infrequent grade 3–4 events, most often stomatitis (~14%)[11]. A number of additional trials in breast cancer are ongoing[12–14].

https://doi.org/10.1038/s41591-024-03266-2

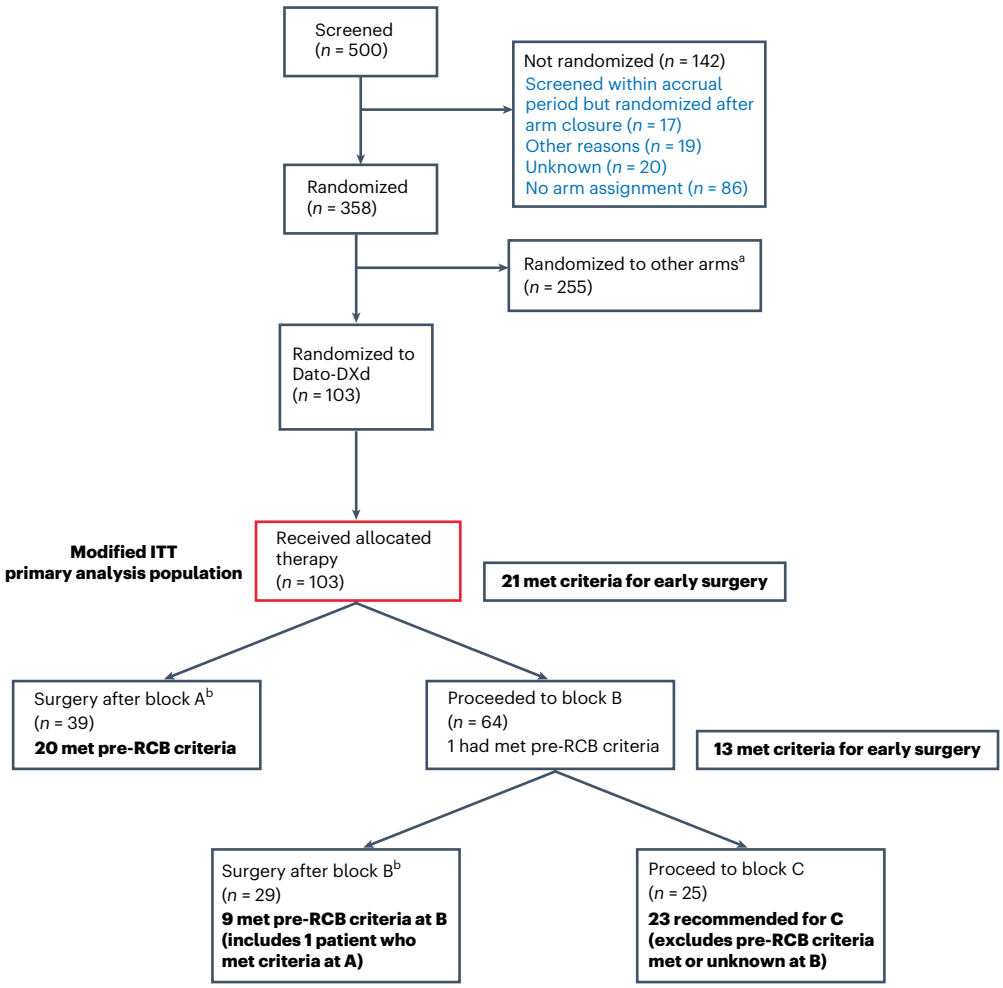


Fig. 2 | CONSORT diagram. The enrollment period was defined as date of first screening consent from arm to date of arm closure to randomization (27 June 2022 to 1 September 2023). The number of patients who adhered to the protocol-specified treatment recommendations for early de-escalation to surgery were also included. [a]Patients randomized to other arms included (1) HR^+HER2^- MammaPrint low risk or clinically node negative MammaPrint high1 class patients randomized to the I-SPY endocrine optimization pilot ($n = 61$), (2) $HER2^+$ patients not eligible to receive Dato-DXd + durvalumab ($n = 46$) and (3) $HER2^-$ patients randomized to other open arms including the Dato-DXd + durvalumab combination ($n = 148$). [b]Includes patients who were assigned non-pCR (at block A or block B, respectively).

The first arm of I-SPY2.2, Dato-DXd, enrolled patients at 44 clinical sites around the continental United States. Here, we describe the results of this arm of the trial and illustrate some of the important lessons learned from this complex and novel patient-centered trial design.

Results

Patients

Between 27 June 2022 and 1 September 2023, 500 patients were screened for the I-SPY2.2 trial (Fig. 2). Of the 358 participants randomized to the trial, 103 were randomized to receive Dato-DXd in the first treatment block of I-SPY2.2 (block A). Enrollment to the arm was halted when maximum accrual of 100 patients was reached, and patients already in the screening process at the time of arm closure were allowed to proceed to randomization and treatment. These 103 patients form the modified intent-to-treat (ITT) population used for the primary analyses. Baseline characteristics of the Dato-DXd population are presented in Table 1. The median age at screening was 46 (range 28–78) years. Most patients were White (58.3%), 11.7% were Hispanic/Latino, 10.7% were Black, 8.7% were Asian and 1% were Indigenous American/Alaska Native.

All patients eligible and randomized to the Dato-DXd arm had HER2-negative disease; 53 (51.5%) had hormone-positive disease, and 50 (48.5%) hormone-negative (triple negative). The most common of the RPS was $HER2^-Immune^+$, accounting for 46 (44.7%) of the 103 participants, followed by $HR^+HER2^-Immune^-DRD^-$ ($n = 36$, 35%), $HR^-HER2^-Immune^-DRD^-$ ($n = 11$, 10.7%) and $HER2^-Immune^-DRD^+$ ($n = 10$, 9.7%). Figure 3a shows how the standard HR/HER2 subtypes map to RPS in this arm. Of note, 17 (of 53) HR^+HER2^- patients were $HER2^-Immune^+$ and were assigned the same block B treatment (paclitaxel + carboplatin + pembrolizumab) as their HR^- counterparts.

The study schema is shown in Fig. 1, illustrating how patients were treated in the trial, beginning with block A (assigned to all subtypes). Figure 2 is the CONSORT diagram, which shows the flow of patients from screening to randomization and demonstrates how many patients proceeded through each block of therapy. Supplementary Fig. 1 shows a simplified schema, annotated by how many patients entered and exited each block of therapy. As shown in the CONSORT, at the completion of treatment in block A, 21 patients met pre-RCB criteria, 20 of whom proceeded to early surgery and 1 continued to block B. A total of 39 patients proceeded to surgery. Sixty-four patients proceeded to block B, including 19 who switched treatment from block A early (Supplementary Fig. 1). At completion of block B, 13 patients met pre-RCB criteria, 9 (including 1 who already met criteria at A) of whom proceeded to early surgery and 4 continued to block C. A total of 29 went to early surgery at block B. Thirty-five patients advanced to block C, 23 of whom were

https://doi.org/10.1038/s41591-024-03266-2

Table 1 | Patient demographics in the Dato-DXd arm

	Overall (*n*=103)
Age (years) at screening	
Mean (s.d.)	47.8 (11.6)
Median (min, max)	46.0 (28.0, 78.0)
Ethnicity	
Hispanic/Latino	12 (11.7%)
Not Hispanic/Latino	83 (80.6%)
Missing	8 (7.8%)
Race	
Asian	9 (8.7%)
Black	11 (10.7%)
Indigenous American/Alaska Native	1 (1.0%)
White	60 (58.3%)
Missing	22 (21.4%)
Subtype	
$HER2^-Immune^-DRD^+$	10 (9.7%)
$HER2^-Immune^+$	46 (44.7%)
$HR^-HER2^-Immune^-DRD^-$	11 (10.7%)
$HR^+HER2^-Immune^-DRD^-$	36 (35.0%)
HR	
Positive	53 (51.5%)
Negative	50 (48.5%)

recommended to proceed to block C. Among the patients proceeding to block C, seven switched treatment from block B early.

Efficacy of Dato-DXd alone (block A)

The efficacy of Dato-DXd treatment in block A is shown in Fig. 3 and Supplementary Table 1. In this analysis, we estimated the pCR rate of block A (alone) using a Bayesian covariate-adjusted model with MRI imputation based on pCR and MRI data. Specifically, we used pCR data from patients exiting block A, and assumed non-pCR at block A for patients with positive biopsy or did not achieve a pCR at later blocks. For patients who did not go to surgery after block A and achieved pCR at later blocks, MRI data are used to impute the probability of pCR at block A and inform the estimated pCR rate. Additional details of the analysis are described in Methods. The probability distributions of the pCR rate for each subtype (Fig. 3b–g) were compared with a fixed pCR rate threshold (represented by a dashed vertical line) set on the basis of the expected response rates for higher and lower responding subtypes. For the higher responding subtypes, for example, $HER2^-Immune^+$, the pCR rate threshold was 40%; for subtypes with lower overall pCR rates to standard of care treatment, the threshold was 15%. Agents are said to 'graduate' if the probability that the pCR rates exceeds the subtype-specific threshold is >85%. Dato-DXd did not meet the block A graduation threshold in any of the subtypes.

Efficacy of treatment strategy

A total of 37 pCRs (38.1%) were observed in the arm over the complete treatment strategy, of which 18 occurred after block A (48.7%), 13 after block B (35.1%) and 6 after block C (16.2%) (Supplementary Fig. 1).

Figure 4 and Supplementary Table 2 show the results for the entire treatment strategy (across all blocks) for each subtype. In this analysis, the pCR rate of the treatment strategy is estimated using a Bayesian model informed by the timing of pCR assessment and pCR status. Modeled rates (Fig. 4, solid lines) are compared against subtype-specific 'dynamic controls' constructed from I-SPY2 data. Details on the dynamic control are described in Methods. Briefly, we estimated pCR rates of candidate comparators selected from previously tested I-SPY2 arms containing elements of current standard of care, including the veliparib + carboplatin arm[15], the pembrolizumab arm[1] and the historical control of paclitaxel followed by doxorubicin–cyclophosphamide (AC), using a Bayesian covariate-adjusted model. Weighted posterior distributions of the candidates are then combined into a single posterior distribution to form the dynamic control (gray filled distributions) against which treatment strategies are compared. Modeled treatment strategies graduate if the probability they are superior to the dynamic control, $P(>DC)$, is >85%.

Given the considerable number of patients who went to early surgery without meeting pre-RCB criteria, we also performed a sensitivity analysis for efficacy including only those individuals who adhered to pre-RCB criteria predicting high likelihood of pCR at time of surgery. Modeled rates from the sensitivity analyses are represented as dashed lines in Fig. 4, and the numbers of patients included in this analysis are highlighted in bold in the CONSORT diagram in Fig. 2.

In the $HR^-HER2^-Immune^-DRD^-$ subtype, a small subtype, 4 of 11 patients (36%) achieved a pCR. Three of the four pCRs (75%) occurred after block A alone (Supplementary Table 3). The modeled pCR rate for the treatment strategy (across all blocks) was 41% (confidence interval (CI) 16–66%). Compared with the dynamic control with a mean estimated pCR rate of 16% (CI 6–26%), the treatment strategy met criteria for graduation with a $P(>DC)$ of 0.97 within this subtype (Fig. 4b and Supplementary Table 2). On sensitivity analysis that included only patients who met criteria to proceed to surgery ($n = 4$) as recommended by the trial's pre-RCB algorithm, the modeled pCR rate was higher at 53% (CI 16–90%) (Fig. 4b).

The graduation threshold for treatment strategies (across all blocks) were not met in any other subtypes. The highest pCR rate was seen in the $HER2^-Immune^+$ subtype, where 27 of 46 patients (59%) achieved pCR. Twelve of the 27 (44%) pCRs were achieved following block A (Supplementary Table 3). In this subtype, the modeled pCR for the treatment strategy was 58% (CI 44–72%), compared with the dynamic control mean pCR rate of 78% (CI 66–90%), with a probability of being superior to the dynamic control ($P(>DC)$) of 0.02. The sensitivity analysis included only patients who adhered to pre-RCB recommendations to proceed to surgery ($n = 25$) and had a higher pCR rate of 67% (CI 49–85%), which was not statistically different from the dynamic control ($P(>DC) = 0.15$) (Fig. 4c and Supplementary Table 2).

In the $HER2^-Immune^-DRD^+$ subtype, a small subset, the modeled pCR for the Dato-DXd/paclitaxel + carboplatin + pembrolizumab/AC + pembrolizumab (D/TCPe/ACPe) treatment strategy was 37% (CI 10–64%) compared with the dynamic control mean pCR rate of 72% (CI 52–92%), with $P(>DC) = 0.03$. The sensitivity analysis indicated that adherence to pre-RCB recommendations resulted in a higher modeled pCR rate of 50% (CI 15–85%, $n = 6$), which was not statistically different from the dynamic control ($P(>DC) = 0.16$) (Fig. 4d and Supplementary Table 2).

The rates of pCR were lowest in the $HR^+HER2^-Immune^-DRD^-$ subtype (3/36 patients, 8%). One pCR occurred after each of block A, block B and block C treatment, respectively (Supplementary Table 3). The modeled pCR rate for the complete treatment strategy was 12% (CI 2–22%) compared with the dynamic control mean pCR rate of 15% (CI 6–24%, $P(>DC) = 0.33$). The sensitivity analysis of pre-RCB adherents resulted in a similar modeled pCR rate (Fig. 4a and Supplementary Table 2).

Considering the immunohistochemical subtype, HR^-HER2^- or triple-negative breast cancer in this arm, 25 of 46 patients (54%) achieved a pCR, with 15 of the 25 observed pCRs (60%) achieved after block A (Supplementary Table 3). The modeled pCR for the complete treatment strategy was 52% (CI 38–66%) compared with the dynamic control mean pCR rate of 65% (CI 56–74%) ($P(>DC) = 0.06$).

https://doi.org/10.1038/s41591-024-03266-2

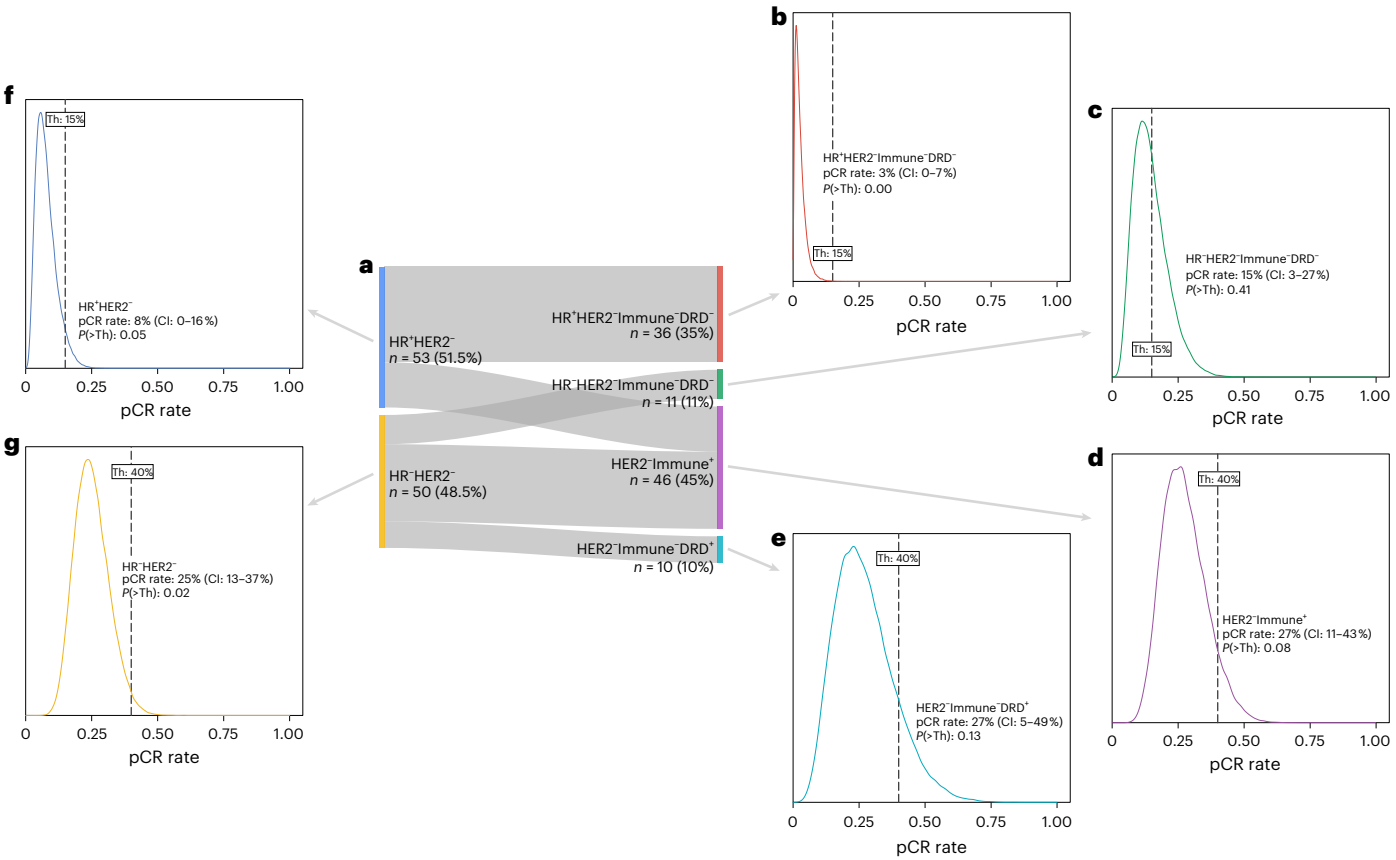


Fig. 3 | Efficacy of Dato-DXd treatment in block A. a, A Sankey diagram illustrating how the RPS relate to standard HR/HER2 subtypes in the trial, with number of patients of each subtype. **b**–**g**, The posterior probability distributions of pCR rate following treatment in block A with Dato-DXd in each of the four RPS subtypes and HER2$^-$ subtypes: HR$^+$HER2$^-$Immune$^-$DRD$^-$ (**b**), HR$^-$HER2$^-$Immune$^-$DRD$^-$ (**c**), HER2$^-$Immune$^+$ (**d**), HER2$^-$Immune$^-$DRD$^+$ (**e**), HR$^+$HER2$^-$ (**f**) and HR$^-$HER2$^-$ (**g**). The mean pCR rate with 95% CI is shown, along with the probability (P(>Th)) that the agent is greater than a preset pCR rate threshold set for a specific subtype. If the P(>Th) is greater than 85% for the threshold (dashed vertical line), the agent is said to 'graduate'.

The sensitivity analysis including only patients who met criteria to proceed to surgery (n = 24) yielded a higher modeled pCR rate of 66% (CI 48–84%) for the treatment strategy, which was not statistically different from the dynamic control (P(>DC) = 0.55) (Fig. 4e and Supplementary Table 2).

Additional granular information on the degree of residual cancer is provided by RCB, a secondary endpoint of the trial. The RCB class distribution (assessed at the time of surgery across all blocks) by subtype are shown in Supplementary Fig. 2. Notably, an additional two (20%) and five (11%) of HR$^-$HER2$^-$Immune$^-$DRD$^-$ and HER2$^-$Immune$^+$ patients had RCB-I, for a total of 60% and 70%, respectively, achieving no or minimal residual disease within these subtypes for the treatment strategy D/TCPe/ACPe.

Note that the planned secondary endpoint of event-free survival is not reported in this paper.

Safety and toxicity

Dato-DXd administered in block A was generally well tolerated by participants. The incidence of adverse events that were experienced by at least 20% of patients during block A is shown in Fig. 5a. The most common adverse events due to Dato-DXd treatment were nausea (91%: 67% grade 1 and 21% grade 2), fatigue (82%: 69% grade 1 and 11% grade 2), rash (75%: 62% grade 1 and 12% grade 2), stomatitis (69%: 52% grade 1 and 15% grade 2) and alopecia (65%: 45% grade 1 and 20% grade 2). Ocular events were noted in 38% of patients; all were grade 1 or 2 (35% grade 1 and 3% grade 2). These included keratitis, eye pain, vision blurred, eye irritation, photophobia, eye pruritus, dry eye, photopsia, conjunctivitis, eye infection, eyelid irritation, corneal ulcer, uveitis, corneal opacity and eye disorder. Notably, there were three cases of grade 3 nausea (3%), two cases of grade 3 fatigue (2%), two cases of grade 3 stomatitis (2%) and one case of grade 3 rash (1%). No grade 4 or 5 events were noted in block A. In block A, no patients discontinued Dato-DXd due to adverse events, although adverse events led to holds in six patients (5.8%), and nine patients (8.7%) underwent Dato-DXd dose modifications due to adverse events (Supplementary Table 4).

Adverse events occurring in at least 20% of patients at any point during the treatment strategy are shown in Fig. 5b. Additional adverse events emerged during blocks B (in bold on the list of adverse events on the y axis) and C, including neuropathy (48%: 38% grade 1, 8% grade 2, 1% grade 3 and 1% grade 4) and neutropenia (22%: 5% grade 1, 5% grade 2, 11% grade 3 and 2% grade 4). Stomatitis and rash were reported in 76% and 79% of patients, respectively, most of which remained at grade 1 or 2 severity.

Reported immune-related adverse events based on the block of onset and exit block for patients off treatment are shown in Fig. 5c. Adrenal insufficiency was seen in four patients, with onset in three patients during block B and one patient during block C. Pneumonitis, colitis, hepatitis, myocarditis, type 1 diabetes mellitus and hyperthyroidism were each seen in one patient, all with onset in block B.

Discussion

One of the primary goals of the I-SPY2.2 design is to get every patient to pCR while minimizing toxicity. However, in the event that a patient goes to the operating room early and the pathology shows residual disease,

https://doi.org/10.1038/s41591-024-03266-2

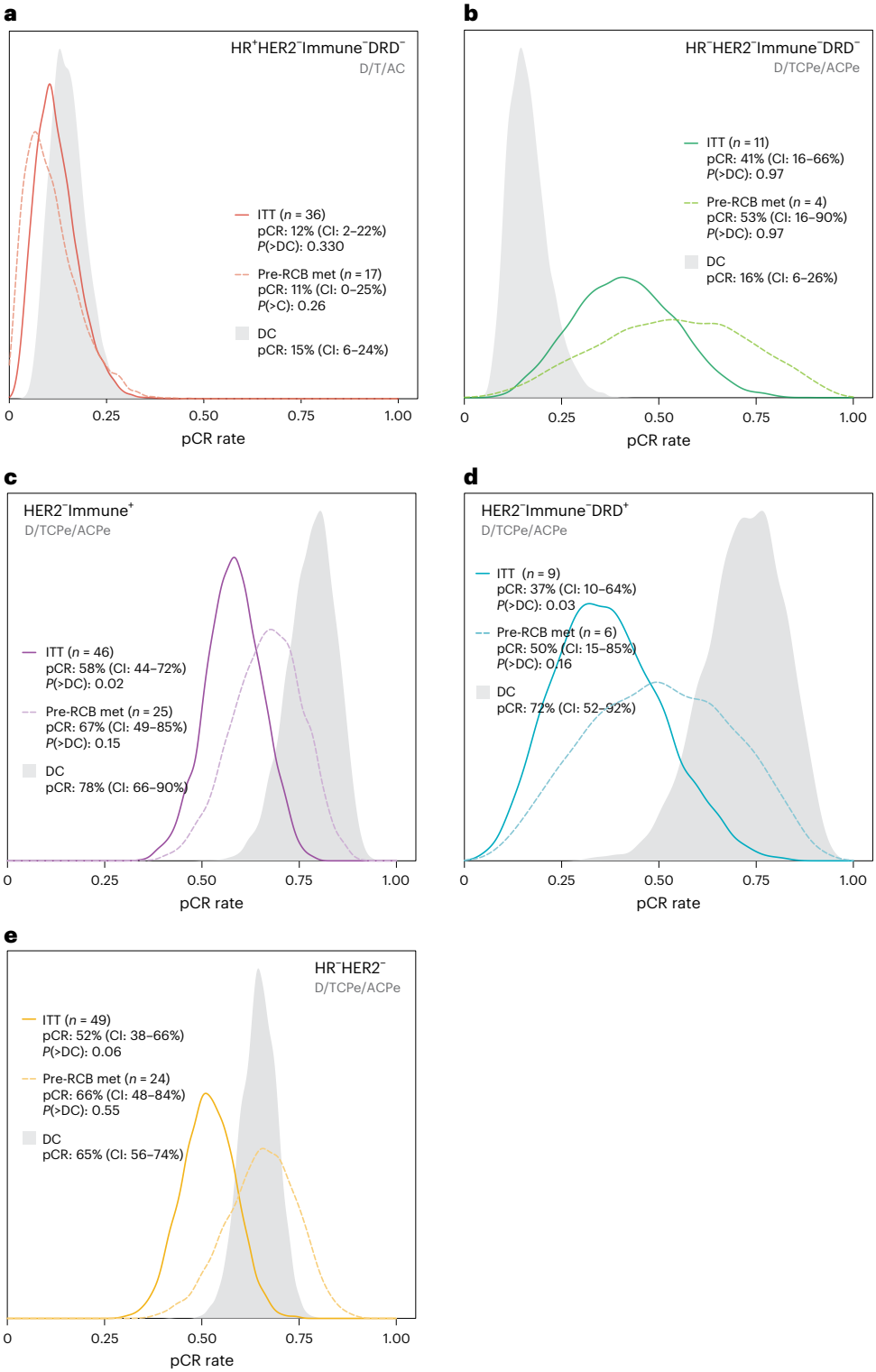


Fig. 4 | Efficacy of treatment strategies spanning blocks A–C. a–e, The posterior pCR rate distributions for each RPS subtype: HR$^+$HER2$^-$Immune$^-$DRD$^-$ (**a**), HR$^-$HER2$^-$Immune$^-$DRD$^-$ (**b**), HER2$^-$Immune$^+$ (**c**), HER2$^-$Immune$^-$DRD$^+$ (**d**) and HR$^-$HER2$^-$ (**e**) and associated treatment strategy (D, Dato-DXd; T, paclitaxel; C, carboplatin; Pe, pembrolizumab) along with the mean pCR rate and 95% CI. The ITT population is the solid line. The dashed line represents a sensitivity analysis that considers only patients who adhered to pre-RCB recommendations following block A and B treatment. *P*(>DC) is the probability that the pCR rate of the treatment strategy is superior to a subtype-specific dynamic control (filled gray distribution) that was predefined using cumulative data from I-SPY2 reflecting current best-in-subtype treatment. The threshold for graduation of a treatment strategy is *P*(>DC) ≥ 0.85. Note that, within the HER2$^-$Immune$^-$DRD$^+$ subtype ($n = 10$), 7 patients went to surgery or received nonprotocol therapy after block A and did not proceed to receive block B treatment. Among the three patients proceeding to block B, two were randomized to TCPe and yielded nine patients who followed the treatment strategy D/TCPe/ACPe shown in **d**.

https://doi.org/10.1038/s41591-024-03266-2

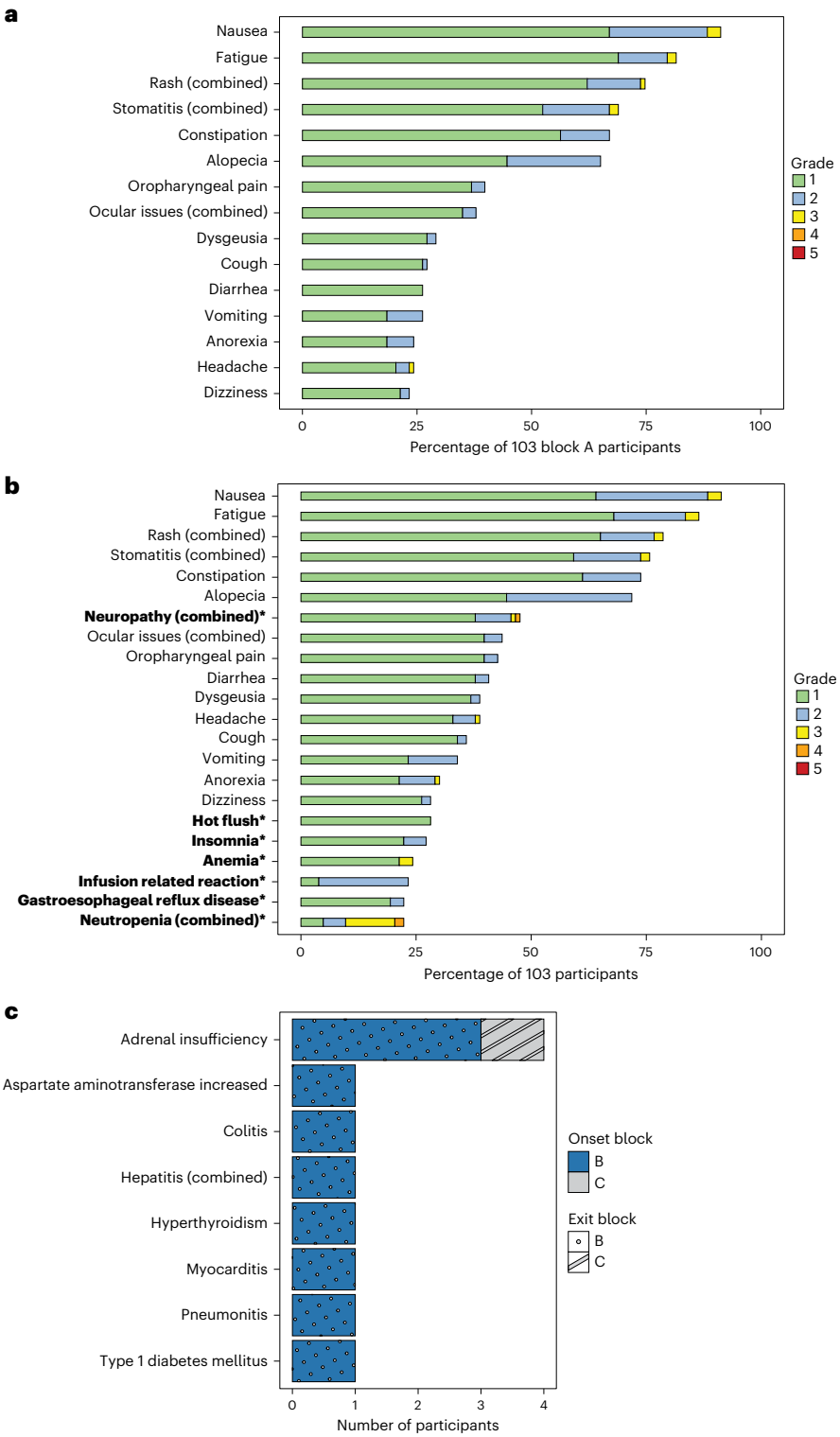


Fig. 5 | Incidence of most frequently observed adverse events. a, The incidence of adverse events of any grade that were experienced by greater than 20% of participants during block A treatment with Dato-DXd ($n = 103$). **b**, The incidence of adverse events of any grade that were experienced by greater than 20% of patients during any block of treatment ($n = 103$). Note that all grade 3/4 adverse events observed in the study are represented here. The adverse events that were associated only with blocks B and C are bolded. **c**, The immune-related adverse events experienced by participants in any treatment block. Note that 36 participants received pembrolizumab + paclitaxel + carboplatin in block B and 16 received AC + pembrolizumab in block C. This does not include the seven patients who withdrew or received off-study therapy (which may include pembrolizumab). Adverse events listed as 'combined' refer to iMedDRA combined terms as follows: stomatitis (combined): stomatitis, oropharyngeal pain, mouth ulceration, mouth injury, oral pain and gingival pain; rash (combined): rash, rash maculo-papular, dermatitis, rash pustular, skin disorder, dermatitis acneiform, rash pruritic, rash erythematous, eczema, urticaria and rash macular; ocular events (combined): keratitis, eye pain, vision blurred, eye irritation, photophobia, eye pruritus, dry eye, photopsia, conjunctivitis, eye infection, eyelid irritation, corneal ulcer, uveitis, corneal opacity and eye disorder; neuropathy (combined): peripheral sensory neuropathy, peripheral motor neuropathy and neuropathy peripheral; neutropenia (combined): febrile neutropenia, neutropenia and neutrophil count decrease.

additional targeted therapy can be administered in the adjuvant setting. In this way, individuals in the trial can try innovative approaches and be safely and optimally managed.

While Dato-DXd monotherapy in the neoadjuvant setting in I-SPY2.2 did not meet the predefined graduation threshold as a single therapy in block A, it performed better than the dynamic control in the HR$^-$HER2$^-$DRD$^-$ subtype when it was considered as part of a treatment strategy (across all blocks). In the block A efficacy analysis, covariate modeling is used to estimate pCR rates as the trial proceeds, which enables us to leverage the full sample size to estimate pCR when there is data missing (as not all patients go to surgery after each block). However, the covariate model tends to underestimate the pCR rates in small subtypes with poor expected response, and this may have resulted in a lower estimated pCR rate (relative to the observed rate) after block A alone in the HR$^-$HER2$^-$Immune$^-$DRD$^-$ subtype. However, the treatment strategy (across all blocks) was clearly active and had a 97% chance of outperforming the dynamic control in this subtype, which is generally expected to have a low pCR rate. Although only 11 patients were in this subset, of the four patients that achieved a pCR, most of the pCRs (three out of four, 75%) occurred after block A alone, and two patients (20%) achieved RCB-1 disease. These results suggest that Dato-DXd is highly active in this subtype and warrants further investigation. We have further optimized the Immune$^-$ subtype for triple-negative cases in I-SPY2 and it is possible that the application of this new classifier could affect the results.

In patients with Immune+ disease, 27 of 46 (59%) achieved a pCR and 25 of 49 (51%) patients with HR$^-$HER2$^-$ disease achieved a pCR by completion of the treatment strategy. The modeled rate of pCR in the Immune$^+$ subtype was 58% (CI 44–72%), which is lower than the dynamic control (78%, CI 66–90%). Because a number of patients went to surgery early despite not meeting pre-RCB criteria, we performed a sensitivity analysis that included only patients who adhered to criteria to proceed to surgery as recommended by the trial's pre-RCB algorithm. In this sensitivity scenario, the modeled pCR rate was higher (67% (CI 49–85%)) in the Immune$^+$ subtype, and 66% (CI 48–84%) in HR$^-$HER2$^-$ disease or triple-negative disease, and not statistically different from the dynamic control. In the study population evaluated in this way, Dato-DXd monotherapy appeared to be an active treatment strategy. Around half of the pCRs in the Immune$^+$ and TNBC subtypes (12/27 (44.4%) and 15/25 (60%), respectively) were achieved after block A.

The sensitivity analysis demonstrated that Dato-DXd did not lower the rate of neoadjuvant response when tested in sequence with standard of care, even in highly responsive subtypes. Importantly, Dato-DXd allowed 18.5% (18/103) of all patients to achieve a pCR after block A without additional therapy, with the associated reduction in toxicity. When we consider the 37 total patients who achieved a pCR (38.1%) across all blocks, almost half of them (18/37, 48.7%) occurred after block A alone, and another 35.1% (13/37) occurred after block B.

The most common adverse events on the Dato-DXd block A of therapy included nausea, rash and fatigue, most of which were of grade 1–2 severity. Alopecia was noted in more than half of patients (65%) during block A. Stomatitis was noted in 69% of patients and ocular events in 38% of patients, despite the trial-mandated use of dexamethasone mouthwash four times daily and lubricating eye drops; but rates of both of these toxicities were mostly of grade 1–2 severity. Stomatitis was more persistent than ocular toxicity throughout the three blocks of therapy, but most events resolved by end of treatment and remained low grade.

As expected, new events emerged on taxane-based block B and anthracycline-based block C of therapy, including neutropenia in 22.3% of patients, with 10.6% with grade 3 and 1.9% with grade 4 neutropenia. Neuropathy emerged in less than half of patients (47.5%), and most events were predominantly grade 1 or 2 in severity. Neither of these events were present in block A.

The pre-RCB algorithm previously described separates patients with a very high chance of having a pCR (in the range of 90%) and those who do not (30–40%). One major limitation of this trial arm was lack of adherence to trial guidance on escalation or de-escalation strategies recommended as a result of pre-RCB assessments. Of the 103 patients enrolled in block A, 39 exited the trial after block A for various reasons, including proceeding to surgery early without pre-RCB prediction of pCR. We are currently investigating the reasons for nonadherence to the recommended treatment strategy, and plan to publish details of this investigation once it is complete. New rules are now in place for real-time monitoring of any cases where surgery is planned for a patient that does not meet the pre-RCB criteria for predicted pCR. Patients from the Dato-DXd arm continue to be followed, and data are also being reviewed on adjuvant therapies received, especially for patients who exited blocks prematurely and did not achieve a pCR. Changes to the protocol and process have improved adherence to the recommendations for early treatment redirection in subsequent arms of I-SPY2.2.

Two additional limitations are the use of a dynamic control that includes only data from I-SPY2 and the small size of some subtypes (including HR$^-$Immune$^-$DRD$^-$ where the agent graduated). I-SPY2.2 uses RPS to incorporate additional biological information predictive of sensitivity to agents targeting immune checkpoints and DRD; as it is relatively recent, it is not yet employed outside of I-SPY2. Thus, our dynamic control is limited to using I-SPY2 data (from ~1,800 patients) and does not include a carboplatin + pembrolizumab + paclitaxel regimen. It should be noted, however, that the estimated pCR rates for the dynamic control of triple-negative patients is 65%, which is similar to the rate observed in the Keynote-522 trial[2].

The Dato-DXd monotherapy arm was one of the first arms on the novel I-SPY2.2 trial design, and it is evident from the sensitivity analysis that, for patients who adhered to the pre-RCB recommendations and followed the intended treatment strategy, their rates of pCR were noninferior to the dynamic control, and in the HR$^-$HER2$^-$Immune$^-$DRD$^-$ subtype, were superior. This particular subtype, which is a subtype that molecularly would not be predicted to respond robustly to HER2-directed therapy, immunotherapy (such as pembrolizumab) or DNA damage repair-targeting agents (such as platinum-based chemotherapy or poly ADP ribose polymerase inhibitors), is a challenging subtype to treat and could potentially be a subtype of breast cancer that benefits most from more novel targeted therapies. Dato-DXd is an ADC targeting TROP2, a transmembrane protein broadly expressed and associated with poor prognosis in HER2$^-$ breast cancer[9], and previous published work has detailed its potent drug–antibody ratio and limited off-target adverse effects[8]. Patients in this trial who were able to achieve pCR after just four cycles of Dato-DXd were spared the toxicity of 24 weeks of additional chemo-immunotherapy, which arguably may not have improved their outcomes further. Neoadjuvant trials with eight cycles of this agent are ongoing, as are trials in the postneoadjuvant setting in patients with non-pCR to traditional chemo-immunotherapy.

We were able to safely evaluate Dato-DXd in early breast cancer and learn more about the ability of four cycles of this agent to achieve pCR without adversely affecting patient outcomes. Further, the efficacy of a combination strategy of Dato-DXd and durvalumab, which enrolled in parallel with this arm, showed promising results that are reported separately in ref. 16. Future studies should incorporate the molecular profile of patients so that we can prospectively identify those most likely to respond to Dato-DXd.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-024-03266-2.

References

1. Nanda, R. et al. Effect of pembrolizumab plus neoadjuvant chemotherapy on pathologic complete response in women with early-stage breast cancer. *JAMA Oncol.* **6**, 676–684 (2020).
2. Schmid, P. et al. Pembrolizumab for early triple-negative breast cancer. *N. Engl. J. Med.* **382**, 810–821 (2020).
3. I-SPY2 Trial Consortium. Association of event-free and distant recurrence–free survival with individual-level pathologic complete response in neoadjuvant treatment of stages 2 and 3 breast cancer. *JAMA Oncol.* **6**, 1355–1362 (2020).
4. Li, W. et al. Abstract P4-02-10: MRI models by response predictive subtype for predicting pathologic complete response. *Cancer Res.* **83**, P4-02-10 (2023).
5. Wolf, D. M. et al. Redefining breast cancer subtypes to guide treatment prioritization and maximize response: Predictive biomarkers across 10 cancer therapies. *Cancer Cell* **40**, 609–623.e6 (2022).
6. Onishi, N. et al. Abstract P3-03-01: functional tumor volume at 3 and 6 week MRI as an indicator of patients with inferior outcome after neoadjuvant chemotherapy. *Cancer Res.* **82**, P3-03-01 (2022).
7. Onishi, N. et al. Prospective performance of an MRI algorithm for early re-direction of breast cancer neoadjuvant treatment. In *Proc. 32nd Annual Scientific Meeting and Exhibition of the International Society for Magnetic Resonance in Medicine* (International Society for Magnetic Resonance in Medicine, 2024).
8. Okajima, D. et al. Datopotamab deruxtecan (Dato-DXd), a novel TROP2-directed antibody–drug conjugate, demonstrates potent antitumor activity by efficient drug delivery to tumor cells. *Mol. Cancer Ther.* **20**, 2329–2340 (2021).
9. Sakach, E., Sacks, R. & Kalinsky, K. Trop-2 as a therapeutic target in breast cancer. *Cancers* **14**, 5936 (2022).
10. Bardia, A. et al. Datopotamab deruxtecan in advanced or metastatic HR+/HER2– and triple-negative breast cancer: results from the phase I TROPION-PanTumor01 study. *J. Clin. Oncol.* **42**, 2281–2294 (2024).
11. Gadaleta-Caldarola, G. et al. Safety evaluation of datopotamab deruxtecan for triple-negative breast cancer: a meta-analysis. *Cancer Treat. Res. Commun.* **37**, 100775 (2023).
12. Dent, R. A. et al. TROPION-Breast02: datopotamab deruxtecan for locally recurrent inoperable or metastatic triple-negative breast cancer. *Future Oncol.* **19**, 2349–2359 (2023).
13. Bardia, A. et al. TROPION-Breast03: a randomized phase III global trial of datopotamab deruxtecan ± durvalumab in patients with triple-negative breast cancer and residual invasive disease at surgical resection after neoadjuvant therapy. *Ther. Adv. Med. Oncol.* **16**, 17588359241248336 (2024).
14. Bardia, A. et al. TROPION-Breast01: datopotamab deruxtecan vs chemotherapy in pre-treated inoperable or metastatic HR+/HER2– breast cancer. *Futur. Oncol.* **20**, 423–436 (2024).
15. Rugo, H. S. et al. Adaptive randomization of veliparib–carboplatin treatment in breast cancer. *N. Engl. J. Med.* **375**, 23–34 (2016).
16. Shatsky, R. A. et al. Datopotamab–deruxtecan plus durvalumab in early-stage breast cancer: the sequential multiple assignment randomized I-SPY2.2 phase 2 trial. *Nat. Med.* https://doi.org/10.1038/s41591-024-03267-1 (2024).

Katia Khoury[1,29], Jane L. Meisel[2,29], Christina Yau[3], Hope S. Rugo[3], Rita Nanda[4], Marie Davidian[5], Butch Tsiatis[5], A. Jo Chien[3], Anne M. Wallace[6], Mili Arora[7], Mariya Rozenblit[8], Dawn L. Hershman[9], Alexandra Zimmer[10], Amy S. Clark[11], Heather Beckwith[12], Anthony D. Elias[13], Erica Stringer-Reasor[1], Judy C. Boughey[14], Chaitali Nangia[15], Christos Vaklavas[16], Coral Omene[17,18], Kathy S. Albain[19], Kevin M. Kalinsky[2], Claudine Isaacs[20], Jennifer Tseng[21], Evanthia T. Roussos Torres[22], Brittani Thomas[23], Alexandra Thomas[24], Amy Sanford[25], Ronald Balassanian[3], Cheryl Ewing[3], Kay Yeung[6], Candice Sauder[7], Tara Sanft[8], Lajos Pusztai[8], Meghna S. Trivedi[9], Ashton Outhaythip[10], Wen Li[3], Natsuko Onishi[3], Adam L. Asare[3,26], Philip Beineke[26], Peter Norwood[26], Lamorna Brown-Swigart[3], Gillian L. Hirst[3], Jeffrey B. Matthews[3], Brian Moore[24], W. Fraser Symmans[27], Elissa Price[3], Carolyn Beedle[3], Jane Perlmutter[28], Paula Pohlmann[27], Rebecca A. Shatsky[6], Angela DeMichele[11], Douglas Yee[12], Laura J. van 't Veer[3], Nola M. Hylton[3] & Laura J. Esserman[3] ✉

[1]University of Alabama at Birmingham, Birmingham, AL, USA. [2]Emory University, Atlanta, GA, USA. [3]University of California San Francisco, San Francisco, CA, USA. [4]University of Chicago, Chicago, IL, USA. [5]North Carolina State University, Raleigh, NC, USA. [6]University of California San Diego, San Diego, CA, USA. [7]University of California Davis, Davis, CA, USA. [8]Yale University, New Haven, CT, USA. [9]Columbia University, New York, NY, USA. [10]Oregon Health Sciences University, Portland, OR, USA. [11]University of Pennsylvania, Philadelphia, PA, USA. [12]University of Minnesota, Minneapolis, MN, USA. [13]University of Colorado Denver, Denver, CO, USA. [14]The Mayo Clinic, Rochester, MN, USA. [15]HOAG Family Cancer Institute, Newport Beach, CA, USA. [16]Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA. [17]Cooperman Barnabas Medical Center, New Brunswick, NJ, USA. [18]Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA. [19]Stritch School of Medicine, Loyola University Chicago, Chicago, IL, USA. [20]Lombardi Comprehensive Cancer Center, Georgetown University, Washington DC, USA. [21]City of Hope Orange County Lennar Foundation Cancer Center, Orange County, CA, USA. [22]University of Southern California, Los Angeles, CA, USA. [23]Sparrow Health System, Lansing, MI, USA. [24]Wake Forest University, Winston–Salem, NC, USA. [25]Sanford Health, Sioux Falls, SD, USA. [26]Quantum Leap Healthcare Collaborative, San Francisco, CA, USA. [27]University of Texas MD Anderson Cancer Center, Houston, TX, USA. [28]The Gemini Group, Ann Arbor, MI, USA. [29]These authors contributed equally: Katia Khoury, Jane L. Meisel.
✉e-mail: laura.esserman@ucsf.edu

Methods

Study design

I-SPY2.2, targeting early breast cancer, is a hybrid design consisting of a sequential multiple assignment randomized trial[17] that integrates a phase 2 adaptive platform approach for early breast cancer at high risk of recurrence (ClinicalTrials.gov identifier NCT01042379). The design enables the development of novel agents and combination treatments, while permitting treatment redirection within the trial, based upon treatment response assessed by MRI. As shown in Fig. 1, the trial consists of three 'blocks' of treatment:

Treatment block A. Treatment block A uses the platform approach to evaluate multiple novel agents in parallel. The master protocol of I-SPY2.2 allows agents to leave or enter block A investigation using protocol amendments. Block A agents are administered as single agents (or combinations) without paclitaxel. MRI is used to assess treatment response longitudinally by assessing change in functional tumor volume (FTV) from pretreatment baseline and predict the probability of pCR[18]. Following completion of a 12-week regimen, participants with MRI-predicted pCR probability above prespecified thresholds whose core biopsy at 12 weeks also show no residual cancer at the primary site, are offered the ability to proceed directly to surgery, forgoing treatment in the remaining two treatment blocks. Participants not meeting the combined MRI and biopsy criteria then proceed to treatment block B.

Treatment block B. Block B offers participants treatment with a standard-of-care regimen predicted to yield the highest rate of pCR based on the RPS classification developed as part of I-SPY2 (ref. 5). Six RPS phenotypes are used based upon HR, HER2, immune-responsive, DRD and intrinsic luminal-ness biomarker signatures. In the current arm, only HER2$^-$ subtypes were eligible for randomization: HR$^+$Immune$^-$DRD$^-$, HR$^-$Immune$^-$DRD$^-$, Immune$^+$ and Immune$^-$DRD$^+$. Regimens in block B typically incorporate paclitaxel (Supplementary Table 5). As with block A, participants whose MRI-predicted pCR probability is above threshold after a 12 week course of treatment and who had no residual disease found in a post-treatment core biopsy may proceed directly to surgery, while those not meeting the MRI and biopsy combined criteria are offered rescue therapy in block C.

Treatment block C. In block C, participants are assigned to receive a minimum of chemotherapy with standard of care doxorubicin/cyclophosphamide over an 8 or 12 week treatment cycle.

Early treatment switching. In I-SPY2, we set a threshold below which participants, following 6 weeks of treatment, were predicted to have moderate to extensive residual disease (RCB 2/3) remaining at 12 weeks. Therefore, to minimize toxicity of ineffective treatment, in blocks A and B, participants with <30% FTV reduction from baseline at 3 weeks and subsequently showing <65% FTV reduction from baseline at 6 weeks may forgo the remaining treatment cycles in that block and proceed to the next treatment block[6].

Endpoint. The primary endpoint of I-SPY2.2 is pCR, which is assessed using the residual cancer burden method[19,20]. Efficacy is evaluated within several clinical signatures, including traditional HR/HER2 subtypes as well as in each of the RPS classification for each novel agent following completion of block A treatment and for each unique sequence of treatments received across blocks A, B and C.

Safety is assessed using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v5.0 (ref. 21). This is augmented by patient-reported adverse events using the patient-reported outcomes version of CTCAE (PRO-CTCAE)[22], the PROMIS[23] quality of life questionnaire and the FACT-GP5 (ref. 24) single item question about the impact of adverse events.

Secondary endpoints include RCB score and class[19], 3- and 5-year event-free survival and distant relapse-free survival. All participants were followed for long-term outcomes.

Dynamic control. As clinical standards of breast cancer care evolve over time, and the I-SPY trial is continuously open and enrolling, we realize our comparator (control) arm needs to update as well to appropriately reflect these ongoing clinical advances. To address this, in I-SPY2.2, we utilize a 'dynamic control' in which we identified a set of candidate comparator arms for each subtype from previously tested I-SPY2 agents that include elements of newer clinical standards (for example, checkpoint inhibitors, carboplatin, dual HER2-targeting in combination with paclitaxel followed by AC). Candidate comparators forming the dynamic control for each subtype are presented in Supplementary Table 6.

Using a Bayesian covariate-adjusted model (described in detail in 'Statistical analysis' section) and data from 1,818 I-SPY2 patients enrolled between March 2010 and April 2022 across 20 arms, we estimated the posterior distribution of pCR rates of each candidate comparator arm. The dynamic control is defined as the weighted posterior distribution of each candidate comparator arm, where the weight is the probability that it is the best arm among the candidate comparator arms within the dynamic control. This weighting forms a single posterior distribution for the dynamic control against which I-SPY2.2 treatment strategies are compared.

Randomization. When Dato-DXd was open to enrollment, participants were randomized with equal probability to open arms at block A for their subtype. The number of open block A arms in each subtype varied over the course of Dato-DXd enrollment. The Dato-DXd plus durvalumab arm was also opened to enrollment during this period and patients were randomized to the Dato-DXd alone and Dato-DXd + durvalumab arms at equal probability. In block B, participants with the Immune$^-$DRD$^+$ subtype were randomized 1:1 to receive one of two block B treatment courses: paclitaxel plus carboplatin or paclitaxel plus carboplatin and pembrolizumab. For all other subtypes, block B treatment was assigned to a single available course of treatment for this block.

Efficacy assessment. Like I-SPY2, I-SPY2.2 is a signal-finding trial that aims to rapidly identify agents or treatment sequences that are likely to be successful in phase 3 trials in responsive subtypes. I-SPY2.2 assesses efficacy in two different ways: (1) evaluating the efficacy of experimental agents/combinations immediately following completion of treatment in block A and (2) evaluating the efficacy of total treatment administered to individuals, in the form of each unique treatment strategy that occurs across the different agents used in blocks A–C.

In assessing efficacy of the experimental agent in block A alone, agents deemed good candidates for further evaluation are said to 'graduate' if there is an 85% probability that their pCR rate exceeds a predetermined fixed pCR rate threshold. These thresholds are selected to reflect a level of activity considered clinically interesting for a novel regimen when given alone without standard chemo/targeted therapies. There are two thresholds, one for subtypes with high response rates to previously evaluated I-SPY2 agents, set at 40%, and one for subtypes with lower response rates, set at 15%.

In assessing efficacy across all blocks, graduation is determined by a comparison to a dynamic control based on the pCR rates expected if participants had been given standard of care or the previous best-in-class sequence from I-SPY2. An agent graduates if the probability that the estimated pCR rate of the sequence is superior to the dynamic control is greater than or equal to 85% in any given clinical signature (subtype).

Both efficacy analyses are modified ITT analyses where all participants receiving on-study therapy are considered evaluable. Participants who withdrew before surgery or receive nonprotocol therapy

before surgery are classified as non-pCR, and their block of exit is classified as the last block of on-study therapy they received.

Efficacy assessment begins after 20 patients have been enrolled onto an arm; pCR rates are monitored by the Data Safety Monitoring Board as enrollment accumulates. For agents open to $HER2^-$ (or $HER2^+$) patients, a maximum sample size of 100 was preset based on simulations of the trial operating characteristics for the efficacy analysis of block A alone. An arm exits the trial upon reaching maximum accrual. However, as in I-SPY2, if the graduation threshold in the efficacy analysis across all blocks is reached within at least one subtype before maximum accrual, the arm is a candidate for exiting the trial early.

Eligibility

I-SPY2.2 is open to women and men aged ≥18 years with anatomic stage II or III breast cancer whose primary tumors are ≥2.5 cm by clinical examination or ≥2.0 cm by imaging. Participants must have an Eastern Cooperative Oncology Group performance status 0 or 1 (ref. 25). Those with MammaPrint low-risk HR^+, $HER2^-$ were excluded from participation due to their low risk of recurrence[26]. All participants signed written informed consent before screening and again after randomization; no compensation was provided for participation in the study. Only $HER2^-$ participants were eligible for randomization to the Dato-DXd arm. Additional eligibility criteria are available in the study protocol.

Trial oversight

I-SPY2.2 was designed and conducted by I-SPY Consortium members, patient advocates and investigators. AstraZeneca provided funds and study drugs (Dato-DXd) but played no role in the study design, collection/analysis of data or in manuscript preparation. Pembrolizumab was purchased for HR^+ patients who were classified as having the immune subtype. The study was approved by Wake Forest School of Medicine, who acted as the central institutional review board (IRB) and a Data Safety Monitoring Board met monthly to review patient safety and study progress. The I-SPY2.2 trial complies with all local and national regulations regarding the use of human study participants and was conducted in accordance to the criteria set by the Declaration of Helsinki. The authors of this manuscript vouch for the accuracy and completeness of the data reported.

Treatments

Participants in the DATO arm received intravenous 6 mg kg^{-1} Dato-DXd (D) on day 1 of each 3 week cycle for up to four cycles in block A.

Participants in block B were prescribed one of three 12 week regimens depending on their RPS classification: paclitaxel (T) only (for $HR^+Immune^-DRD^-$), paclitaxel (T) plus concurrent carboplatin (C) (for $HER2^-Immune^-DRD^+$), or concurrent pembrolizumab (Pe), paclitaxel and carboplatin (for $HR^-HER2^-Immune^-DRD^-$, $HER2^-Immune^+$ and $HER2^-Immune^-DRD^+$). Pembrolizumab was administered intravenously at 200 mg every 3 weeks for four cycles. The target dose of intravenous carboplatin was 1.5 weekly (area under the curve). Paclitaxel was administered intravenously at 80 mg m^{-2} weekly for 12 weeks. For the first paclitaxel infusion, 20 mg dexamethasone was administered; if no infusion reaction was noted, dexamethasone was reduced to 10 mg then discontinued if no further reactions occurred. The order of administration in block B was pembrolizumab, then paclitaxel and then carboplatin. All participants reaching block C were prescribed four cycles of 60 mg m^{-2} intravenous doxorubicin (A) and 600 mg m^{-2} cyclophosphamide (C), administered every 2 or 3 weeks (dose dense or standard AC). For certain subtypes, block C therapy also included four cycles of 200 mg intravenous pembrolizumab every 3 weeks over a total of 12 weeks (Supplementary Table 5).

This created the following sequences of treatment administered by RPS: D/T/AC in $HR^+HER2^-Immune^-DRD^-$; D/TCPe/ACPe in $HR^-HER2^-Immune^-DRD^-$; D/TCPe/ACPe in $HER2^-Immune^+$; and D/TCPe/ACPe and D/TC/ACPe in $HER2^-Immune^-DRD^+$. The $HER2^-Immune^-DRD^+$ subtype was small, and several patients randomized to TC in block B received Pe off study protocol. The efficacy of both strategies (D/TCPe/ACPe and D/TC/ACPe) are presented for completeness in Supplementary Table 2, but only the D/TCPe/ACPe strategy is shown in Fig. 4 and considered in additional analyses (for example, timing of pCR and RCB class distribution). Of note, the above RPS-based treatment assignments result in 49/50 HR^-HER2^- patients (across RPS) receiving treatments consistent with strategy D/TCPe/ACPe (enabling efficacy evaluation of across all blocks for this subtype).

Dose reductions and toxicity management were specified in the protocol. Adverse events were collected according to the NCI CTCAE version 5.0.

Following completion of neoadjuvant therapy, definitive surgery by lumpectomy or mastectomy was performed at the discretion of the treating surgeon. Sentinel node dissection was allowed in node-negative patients, with axillary node dissection in node-positive patients according to National Comprehensive Cancer Network (NCCN) and local practice guidelines[27]. Adjuvant treatment was not mandated by the trial and was left to the discretion of the treating oncologist.

Assessments

Specimens. Sixteen-gauge core needle biopsies from the primary tumor were collected at screening and at the completion of block A and block B (when applicable) and prepared according to standard I-SPY methods[28]; surgical specimens were prepared in similar fashion. Whole blood for plasma and buffy coat aliquots were collected (when applicable) at screening (baseline), and 3, 6 and 12 weeks following the start of treatment in each of blocks A and B, and again before surgery. All unused portions of specimens are banked for future research.

Serial MRI. MRI were performed at screening (baseline), after 3, 6 and 12 weeks of treatment in each of blocks A and B (when applicable), and before surgery. Another MRI at the discretion of the treating physician could be scheduled at 6 weeks following start of clock C treatment. For each contrast-enhanced examination, FTV and change in FTV from baseline (ΔFTV) were calculated using previous published methods[18].

MammaPrint. MammaPrint classification was used for eligibility assessment and in subtype classification. Using baseline specimens, participants were classified as MammaPrint (Agendia) high risk (MP1) or MammaPrint ultrahigh risk (MP2), using a predefined threshold applied to the 70-gene risk score evaluated on Agilent 44K arrays[26,29]. The threshold used is equivalent to the median cut point of I-SPY1 participants (−0.154 in the original I-SPY1 dataset) who fit the eligibility criteria for I-SPY2.

RPS assignment. RPS are defined by combining clinical biomarkers: HR status and HER2 status and classifications from three expression-based signatures computed from Agilent arrays (Agendia): BluePrint, ImPrint and RePrint. Participants were classified using the 80-gene molecular subtyping test BluePrint into luminal versus basal or HER2 subtypes[5]. The ImPrint signature was developed based on data from the I-SPY2 trial as a predictive biomarker of immunotherapy response as previously described[5] and used to classify patients into $Immune^+$ (that is, predicted sensitive) versus $Immune^-$ (that is, predicted insensitive) subsets. The 60-gene RePrint signature, also developed from I-SPY2 trial data to predict sensitivity to DRD targeting agents, was used to classify patients into DRD^+ (that is, predicted sensitive) versus DRD^- (that is, predicted insensitive) subsets. These five biomarkers are combined to form six phenotypic RPS with differential responsiveness to previously evaluated I-SPY2 agents: $HR^+HER2^-Immune^-DRD^-$, $HR^-HER2^-Immune^-DRD^-$, $HER2^-Immune^+$, $HER2^-Immune^-DRD^+$, $HER2^+$nonluminal, $HER2^+$luminal.

Adverse events—adrenal insufficiency adjudication. Reported adverse events were adjudicated and validated by the I-SPY2 safety chairs (H.S.R. and R.N.) in collaboration with Quantum Leap Healthcare

Collaborative. Immune-related adverse events were validated and adjudicated using published guidelines with confirmation of adrenal insufficiency by two independent endocrinologists[30,31].

Return of results

The central IRB chair set up a process whereby results of investigational diagnostics (conducted under an investigational device exemption) were able to be shared with patients. The IRB reviews and approves the proposed test results to share and the patient friendly report in which the information is returned. Return of results are used for reporting information to patients that affect their care. This includes the RPS used for randomization, the MRI FTV and the recommendations for continuing treatment or escalating to the next block of therapy and at 6 weeks, as well as the pre-RCB report and recommendation for whether they have met the threshold for having a high probability of having a compete response, triggering a recommendation to proceed to surgical resection or, if not meeting criteria, to continue on to the next block of therapy.

Statistical analysis

Efficacy assessment of block A alone. For each block A agent, we leveraged existing I-SPY2 statistical methodologies to model the posterior distribution of pCR rates after block A within each signature (for comparison against a fixed pCR rate threshold) as described below.

The Bayesian model used has two main components: (1) a core covariate-adjusted Bayesian logistic regression and (2) an MRI imputation model.

The core logistic regression model. Probability distributions of pCR rates are estimated using a Bayesian covariate-adjusted logistic model with marker statuses (HR, HER2, Immune, BluePrint luminal and DRD) as covariates for each eligible signature.

Let $y_i \in \{0,1\}$ be the indicator for the pCR response of patient $i (i = 1, \ldots, N)$ at block A. Covariates $x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i} \in \{0,1\}$ represent the HR, HER2, Immune, DRD and BluePrint statuses of patient i (with 1 indicating positive and 0 negative for HR, HER2, Immune and DRD, and 'HER2orBasal' and 'Luminal' for BluePrint). Label A_i is the treatment arm assigned to patient i. The full model is

$$y_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + +\beta_4 x_{4i} + \beta_5 x_{5i} + \theta_{A_i} + \gamma_{1,A_i} x_{1i} + \gamma_{2,A_i} x_{2i} + \gamma_{3,A_i} x_{3i} + \gamma_{4,A_i} x_{4i} + \gamma_{5,A_i} x_{5i}$$

The model's components are as follows:

- The $\beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_4 + \beta_5 x_5$ terms capture the effect of being in a particular subgroup defined by (HR, HER2, Immune, DRD BluePrint) status.
- The $\gamma_{1,A_i} x_{1i}, \gamma_{2,A_i} x_{2i}, \gamma_{3,A_i} x_{3i}, \gamma_{4,A_i} x_{4i} \gamma_{5,A_i} x_{5i}$ terms are the treatment effects within each of the (HR, HER2, Immune, DRD and BluePrint) subgroups.
- We set $\theta_0 = \gamma_{1,0} = \gamma_{2,0} = \gamma_{3,0} = \gamma_{4,0} = \gamma_{5,0} = 0$ to ensure parameter identifiability.
- The θ_{A_i} term represents the effect of being on a particular treatment arm for all patients.

In addition to the core logistic regression structure of the above terms, we include imputation of the pCR status based on longitudinal MRI data over time. The MRI imputation component of the model is described below.

MRI imputation model. If all the pCR data y_i were available, then this is a standard Bayesian logistic regression model that may be fit using Markov chain Monte Carlo methods. When pCR data for some patients are not available, we use imputation based on the patients for whom pCR status is known. Of note, patients who did not go to surgery at the end of block A but had a positive 12 week biopsy or did not achieve a pCR at later treatment blocks are considered non-pCR (not missing) for analysis. We then use the imputed values in the logistic regression model.

For modeling missing pCR, we employ a simple imputation model that exploit MRI tumor volume measurements during treatment. We fit this separate imputation logistic regression model as described below to assess the probability of achieving pCR based on the available MRI information. The imputation model is used for modeling missing pCR results, while the previous model describes the modeling of the likelihood of pCR given baseline factors and treatment arm.

Define S_0, S_1, S_2 and S_3 to be the tumor volumes at baseline, 3 weeks, 6 weeks and 12 weeks. Then, the tumor volume reduction at visit k is $r_k = \frac{(S_k - S_0)}{S_0}$ for $k = 1, 2, 3$.

Negative values of r_k correspond to a decrease in tumor volume, with $r_k = -1.0$ indicating that the tumor is not detectable by MRI at visit k. To formulate a multiple imputation model for pCR status in cases where pCR is not available, we discretize the range of r_k values into categories, for $m_k = 1, \ldots, 13$. The mapping between r_k and m_k is provided in Supplementary Table 7.

For each time point $k = 1, 2, 3$, we use data from all patients for whom both the pCR response and MRI assessment r_k are available, and we fit this model:

$$\text{logit}(\pi_{k,i}) = \alpha_{k,m_i}$$

where α_{k,m_i} is the parameter corresponding to the category m_i, where $m_i = 1, \ldots, 13$ (Supplementary Table 7), for patient i at time k. We impose the monotonicity constraint $\alpha_{k,l} \leq \alpha_{k,n}$ for $l < n$ to reflect the condition that greater tumor reduction cannot be less likely to lead to a pCR.

At each Markov chain Monte Carlo iteration, if the pCR status for patient j is missing, a mixture distribution is formed with the predicted probability of pCR ($\pi_{K,j}$), where K is the latest time point for which we have a tumor volume measurement for patient j.

Distribution of molecular tumor subclasses. The population incidence rates across the 32 disease subclasses defined by HR, HER2, Immune, DRD and BluePrint are modeled as a categorical distribution (that is, multinomial distribution with 32 outcomes) over the subtypes indexed by {1, 2,..., 32}, with $\Pr(C = c) = \phi_c$. We use a Dirichlet prior

$$(\phi_1, \phi_2, \ldots, \phi_{32}) \sim \text{Dirichlet}\,(0.1, 0.1, \ldots, 0.1),$$

which results in a posterior distribution

$$(\phi_1, \phi_2, \ldots, \phi_{32}) | \text{Data} \sim \text{Dirichlet}\,(n_1 + 0.1, n_2 + 0.1, \ldots, n_{32} + 0.1)$$

where n_c is the number of patients observed in subclass c.

pCR modeling for signatures. The posterior distribution of the pCR rate for signatures is created from the pCR rates by subtypes contained in each signature and weighted by the prevalence parameters as follows.

Let $\pi_C(D)$ be the probability of pCR for treatment arm a in disease subclass C. If C represents $(x_1, x_2, x_3, x_4, x_5)$, where $x_i \in \{0,1\}$ for the (HR, HER2, Immune, DRD and BluePrint) value, then

$$\text{logit}(\pi_C(a)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \theta_a + \gamma_{1,a} x_1 + \gamma_{2,a} x_2 + \gamma_{3,a} x_{3a} + \gamma_{4,a} x_4 + \gamma_{5,a} x_{5a}$$

Let $P_S(a)$ be the probability of pCR for treatment arm a for signature S. Then

$$P_S(a) = \frac{\sum_{C \in S} \pi_C(a) \phi_C}{\sum_{C \in S} \phi_C}$$

(the notation $C \in S$ means subclass C is included in signature S).

Prior distributions. Prior distributions for coefficients in the core logistic model are taken as normally distributed with s.d. of 1. Prior means for the treatment effects coefficients are assumed to be 0. Prior means for the marker coefficients in the core logistic model and for the MRI imputation coefficients were obtained using the set of $n = 1{,}818$ I-SPY2 patients with full pCR. Since only the 3 and 12 week MRI were collected in I-SPY2, the 6 week MRI imputation mean was linearly interpolated. The distribution of coefficients for each MRI time point was taken as an ordered normal distribution with s.d. of 0.2.

Efficacy assessment of block A agent in the context of a treatment strategy (across all blocks). Efficacy assessment after block A alone is myopic in that it does not consider the performance of the block A regimen as part of a treatment strategy. For instance, a block A agent may potentiate the effect of a block B agent to give a highly effective treatment strategy, even though it did not meet our threshold for success based on the efficacy assessment of block A alone. Thus, we will also evaluate efficacy of block A agents in the context of a treatment strategy.

The efficacy analysis of treatment strategies (across all blocks) uses pCR data and timing of surgery (for example, after block A, B or C) to estimate the pCR rate. A treatment strategy (a, b, c) is a 'path': administer treatment a at block A, if the participant proceeds to block B, then administer treatment b, and if the participant proceeds to block C, administer treatment c. There are three ways a participant can be consistent with this treatment strategy: (1) if they receive a at block A and went to surgery after block A, (2) if they receive a at block A and then b at block B and went to surgery after block B and (3) if they receive a, b and c at their respective blocks. The pCR rate of a treatment strategy is estimated using a combination of five Bayesian models (described below) from data of patients whose treatment is consistent with the strategy.

Bayesian model for pCR rate of strategy. As treatment options are subtype specific and we expect pCR rates to depend on the subtype, we independently conduct this algorithm for each subtype. We describe it in general below. Analysis is performed only for five of the six prespecified signatures (each RPS and HR^-HER2^-), as HR^+HER2^- patients did not share any common block B treatment option across response-predictive subtypes.

Let R_A be the indicator the subject elects to receive surgery at block A; similarly define R_B for block B. Next, let Y_A be the indicator that a subject has a pCR at block A given they went to surgery; similarly define Y_B and Y_C. Let A = a and B = b indicate treatments at blocks A and B, respectively.

Consider the following probabilities:

$$\theta_A(a) = P(R_A = 1 \mid A = a)$$

$$\pi_A(a) = P(Y_A = 1 \mid A = a, R_A = 1)$$

$$\theta_B(a,b) = P(R_B = 1 \mid A = a, R_A = 0, B = b)$$

$$\pi_B(a,b) = P(Y_B = 1 \mid A = a, R_A = 0, B = b, R_B = 1)$$

$$\pi_C(a,b,c) = P(Y_C = 1 \mid A = a, R_A = 0, B = b, R_B = 0, C = c)$$

Using these probabilities, we define the probability of a pCR in the trial given treatment strategy (a, b, c) as

$$\mu(a,b,c) = \theta_A(a)\pi_A(a) + \{1-\theta_A(a)\}\theta_B(a,b)\pi_B(a,b) + \{1-\theta_A(a)\}\{1-\theta_B(a,b)\}\pi_C(a,b,c).$$

This posterior distribution is based on the posterior distributions of $\theta_A(a)$, $\pi_A(a)$, $\theta_B(a,b)$, $\pi_B(a,b)$ and $\pi_C(a,b,c)$. For each of these, we assume beta (0.5, 0.5) prior distributions and update the posteriors with the observed data. For example, if 50 people have been assigned treatment A = a and 35 of those people have $R_A = 1$, then the posterior for $\theta_A(a)|\text{Data} \sim \text{Beta}(0.5 + 35, 0.5 + 50 - 35)$.

Of note, different sets of patients contribute to each of the five Bayesian models: participants who receive a at block A contribute to estimation of $\theta_A(a)$ and those participants who also proceeded to surgery at block A contribute to estimating $\pi_A(a)$. Similarly, participants who receive a at block A and b at block B informs $\theta_B(a,b)$ and those who also proceed to surgery at block B contribute to estimating $\pi_B(a,b)$. Participants who receive a at block A, b at block B and c at block C contribute to the estimation of $\pi_C(a,b,c)$.

Efficacy analysis of block A alone was performed using R and STAN via the RSTAN package, which implements the NUTS sampler. Efficacy analysis of treatment strategy was performed using R via the stats package.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

De-identified subject level data and/or clinical specimens are made available to members of the research community upon approval of the I-SPY Data Access and Publications Committee. Details of the application and review process are available at https://www.quantumleaphealth.org/for-investigators/clinicians-proposal-submissions/. I-SPY aims to make complete patient-level clinical datasets available for public access within 6 months of publication, as the data is complex and requires extensive annotation to ensure its usability.

Code availability

The statistical code used in this clinical trial is available to other investigators for approved research purposes. Investigators interested in accessing the code complete an application at https://www.quantumleaphealth.org/for-investigators/clinicians-proposal-submissions/ and include a brief description of the intended use. Access to the code will be granted upon approval of the request, subject to compliance with ethical guidelines and applicable institutional and regulatory requirements.

References

17. Lavori, P. W. & Dawson, R. Introduction to dynamic treatment strategies and sequential multiple assignment randomization. *Clin. Trials* **11**, 393–399 (2014).
18. Li, W. et al. Predicting breast cancer response to neoadjuvant treatment using multi-feature MRI: results from the I-SPY 2 TRIAL. *NPJ Breast Cancer* **6**, 63 (2020).
19. Symmans, W. F. et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J. Clin. Oncol.* **25**, 4414–4422 (2007).
20. Yau, C. et al. Residual cancer burden after neoadjuvant chemotherapy and long-term survival outcomes in breast cancer: a multicentre pooled analysis of 5161 patients. *Lancet Oncol.* **23**, 149–160 (2022).
21. Common terminology criteria for adverse events (CTCAE) protocol development. *CTEP National Cancer Institute* https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm (2021).
22. Basch, E. et al. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J. Natl Cancer Inst.* **106**, dju244 (2014).
23. Jacob, S. et al. Use of PROMIS to capture patient reported outcomes over time for patients on I-SPY2. *J. Clin. Oncol.* **41**, 611 (2023).

24. Pearman, T. P., Beaumont, J. L., Mroczek, D., O'Connor, M. & Cella, D. Validity and usefulness of a single-item measure of patient-reported bother from side effects of cancer therapy. *Cancer* **124**, 991–997 (2018).
25. Oken, M. M. et al. Toxicity and response criteria of the eastern-cooperative-oncology-group. *Am. J. Clin. Oncol.* **5**, 649–655 (1982).
26. Cardoso, F. et al. 70-Gene signature as an aid to treatment decisions in early-stage breast cancer. *N. Engl. J. Med.* **375**, 717–729 (2016).
27. Beltran, P. J. et al. Ganitumab (AMG 479) inhibits IGF-II–dependent ovarian cancer growth and potentiates platinum-based chemotherapy. *Clin. Cancer Res.* **20**, 2947–2958 (2014).
28. Pusztai, L. et al. Durvalumab with olaparib and paclitaxel for high-risk HER2-negative stage II/III breast cancer: results from the adaptively randomized I-SPY2 trial. *Cancer Cell* **39**, 989–998.e5 (2021).
29. Piccart, M. et al. 70-Gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *Lancet Oncol.* **22**, 476–488 (2021).
30. Brahmer, J. R. et al. Society for Immunotherapy of Cancer (SITC) clinical practice guideline on immune checkpoint inhibitor-related adverse events. *J. Immunother. Cancer* **9**, e002435 (2021).
31. Haanen, J. et al. Management of toxicities from immunotherapy: ESMO clinical practice guideline for diagnosis, treatment and follow-up. *Ann. Oncol.* **33**, 1217–1238 (2022).

Acknowledgements

Research reported in this paper was supported by the NCI of the National Institutes of Health (NIH) under award numbers P01CA210961 and U01CA225427 (L.J.E., N.H.). We acknowledge the generous support of the study sponsors and operations management, Quantum Leap Healthcare Collaborative (QLHC, 2013 to present) and the Foundation for the NIH (2010 to 2012; to L.J.E.). We sincerely appreciate the ongoing support for the I-SPY2.2 Trial from the Safeway Foundation, the William K. Bowes, Jr. Foundation, Give Breast Cancer the Boot and QLHC (all to L.J.E.) and the Breast Cancer Research Foundation (to L.J.E. and L.J.v.V.). We thank all the patients who volunteered to participate in I-SPY2. AstraZeneca provided funds and study drugs (datopotamab–deruxtecan and durvalumab). With the exception of QLHC, the funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank D. Wolf and J. Gibbs and the I-SPY patient advocates for all their important contributions to this work.

Author contributions

Arm chaperones: J.L.M. and K.K. Leadership: C.Y., H.S.R., R.N., B.M., J.P., P.P., R.A.S., A.D., D.Y., L.J.v.V., N.M.H. and L.J.E. Study design: C.Y., M.D., B. Tsiatis, A.D., D.Y., L.J.v.V., N.M.H. and L.J.E. Formal analysis: C.Y. and P.N. Enrolled patients: J.L.M., K.K., H.S.R., R.N., A.J.C., A.M.W., M.A., M.R., D.L.H., A.Z., A.S.C., H.B., A.D.E., E.S.-R., J.C.B., C.N., C.V., C.O., K.S.A., K.M.K., C.I., J.T., E.T.R.T., B.T., A.T., A.S., R.B., C.E., K.Y., C.S., T.S., L.P., M.S.T., A.O., R.A.S., A.D., D.Y. and L.J.E. Laboratory studies: L.B.-S., G.L.H. and L.J.v.V. Imaging/pathology: W.L., N.O., W.F.S. and N.M.H. Data management: A.L.A. and P.B. Administration: G.L.H. and J.B.M. First draft: J.L.M., K.K., C.Y., P.N., J.B.M. and L.J.E. Edit and review: J.L.M., K.K., C.Y., H.S.R., R.N., M.D., B. Tsiatis, A.J.C., A.M.W., M.A., M.R., D.L.H., A.Z., A.S.C., H.B., A.D.E., E.S.-R., J.C.B., C.N., C.V., C.O., K.S.A., K.M.K., C.I., J.T., E.T.R.T., B. Thomas, A.T., A.S., R.B., C.E., K.Y., C.S., T.S., L.P., M.S.T., A.O., W.L., N.O., A.L.A., P.B., P.N., L.B.-S., G.L.H., J.B.M., B.M., W.F.S., E.P., C.B., J.P., P.P., R.A.S., A.D., D.Y., L.J.v.V., N.M.H. and L.J.E.

Competing interests

J.L.M. reports institutional research funding from AstraZeneca, Seagen, Sermonix and Olema and advisory and consulting roles with Pfizer, Seagen, Sermonix, Novartis, Stemline, AstraZeneca, Olema, GlaxoSmithKline and GE Healthcare. C.Y. reports institutional research grant from NCI/NIH; salary support and travel reimbursement from QLHC; a United States patent titled 'Breast cancer response prediction subtypes' (no. 18/174,491); and University of California Inventor Share. H.S.R. reports institutional research support from AstraZeneca, Daiichi Sankyo, Inc., F. Hoffmann–La Roche AG/Genentech, Inc., Gilead Sciences, Inc., Lilly; Merck and Co., Novartis Pharmaceuticals Corporation, Pfizer, Stemline Therapeutics, OBI Pharma, Ambrx, Greenwich Pharma; and advisory and consulting roles with Chugai, Puma, Sanofi, Napo and Mylan. R.N. reports research funding from Arvinas, AstraZeneca, BMS, Corcept Therapeutics, Genentech/Roche, Gilead, GSK, Merck, Novartis, OBI Pharma, OncoSec, Pfizer, Relay, Seattle Genetics, Sun Pharma and Taiho and advisory roles with AstraZeneca, BeyondSpring, Daiichi Sankyo, Exact Sciences, Fujifilm, GE, Gilead, Guardant Health, Infinity, iTeos, Merck, Moderna, Novartis, OBI, OncoSec, Pfizer, Sanofi, Seagen and Stemline. M.D. reports research grants from NIH/NCI and NIH/NIA, and contracts from PCORI. A.J.C. reports institutional research funding from Merck, Amgen, Puma, Seagen, Pfizer and Olema and advisory roles with AstraZeneca and Genentech. A.Z. reports institutional research funding from Merck, honoraria for Medscape and participation on Pfizer Advisory Board. A.S.C. reports institutional research funding from Novartis and Lilly. A.D.E. reports support from Scorpion, Infinity and Deciphera. E.S.-R. reports grants from V Foundation, NIH, Susan G. Komen; institutional research funding from GSK, Seagen, Pfizer, Lilly; consulting and honoraria from Novartis, Merck, Seagen, AstraZeneca, Lilly; Cancer Awareness Network Board member and support from ASCO and NCCN. J.C.B. reports institutional research funding from Eli Lilly and SymBioSis, participation on the Data Safety Monitoring Committee of Cairn Surgical and honoraria from PER, PeerView, OncLive, EndoMag and UpToDate. C.V. reports institutional research funding from Pfizer, Seagen, H3 Biomedicine/Eisai, AstraZeneca, CytomX; research funding to previous institution from Genentech, Roche, Pfizer, Incyte, Pharmacyclics, Novartis, TRACON Pharmaceuticals, Innocrin Pharmaceuticals, Zymeworks and H3 Biomedicine; advisory and consulting roles with Guidepoint, Novartis, Seagen, Daiichi Sankyo, AstraZeneca and Cardinal Health; unpaid consulting with Genentech; and participation in non-CME activity with Gilead, AstraZeneca. C.O. reports consulting fees from AstraZeneca, Guardant Health and Jazz Pharmaceuticals. K.S.A. reports institutional research funding from AstraZeneca, Daiichi Sankyo, Seattle Genetics and QLHC; Independent Data and Safety Monitoring committee at Seattle Genetics. K.M.K. reports advisory and consultant roles for Eli Lilly, Pfizer, Novartis, AstraZeneca, Daiichi Sankyo, Puma, 4D Pharma, OncoSec, Immunomedics, Merck, Seagen, Mersana, Menarini Silicon Biosystems, Myovant, Takeda, Biotheranostics, Regor, Gilead, Prelude Therapeutics, RayzeBio, eFFECTOR Therapeutics and Cullinan Oncology; and reports institutional research funding from Genentech/Roche, Novartis, Eli Lilly, AstraZeneca, Daiichi Sankyo and Ascentage. C.I. reports institutional research funding from Tesaro/GSK, Seattle Genetics, Pfizer, AstraZeneca, BMS, Genentech, Novartis and Regeneron; consultancy roles with AstraZeneca, Genentech, Gilead, ION, Merck, Medscape, MJH Holdings, Novartis, Pfizer, Puma and Seagen; and royalties from Wolters Kluwer (UptoDate) and McGraw Hill (Goodman and Gillman). J.T. serves as institutional principal investigator for clinical trial with Intuitive Surgical; editor lead for ABS, CGSO, SCORE, Breast Education Committee Track Leader, ASCO SESAP 19 and Breast Co-Chair, ACS. A.T. owns stock at Johnson and Johnsons, Gilead, Bristol Myers Squibb, reports participation on

Pfizer Advisory Board: AstraZeneca and reports institutional research funding from Merck and Sanofi and royalties from UptoDate. R.B. reports a consultancy role at Genentech and stock ownership at Cerus Corp. K.Y. received research support unrelated to this work and paid to the institution from Pfizer, Gilead, Seagen, Dantari Pharmaceuticals, Treadwell Therapeutics, and Relay Therapeutics; support from American Cancer Society IRG grant no. IRG-19-230-48-IRG, UC San Diego Moores Cancer Center, Specialized Cancer Center support grant NIH/NCI P30CA023100, Curebound Discovery Award (2023, 2024). T.S. reports honoraria from Hologic. L.P. reports institutional research funding from Susan Komen Foundation, Breast Cancer Research Foundation, NCI, Pfizer, AstraZeneca, Menarini/Stemline, Bristol Myers Squibb, Merck and Co.; consulting fees from AstraZeneca, Merck, Novartis, Genentech, Natera, Personalis, Exact Sciences and Stemline/Menarini; patent titled 'Method of measuring residual cancer and predicting patient survival' (no. 7711494); and Data and Safety Monitoring Board member of the DYNASTY Breast02, OPTIMA and PARTNER trials. M.S.T. reports institutional research funding from Lilly, Gilead Sciences, Phoenix Molecular Designs, AstraZeneca, Regeneron, Merck and Novartis. A.L.A., P.B. and P.N. are employees of QLHC. G.L.H. reports institutional research grant from NIH (1R01CA255442). W.F.S. reports shares of IONIS Pharmaceuticals and Eiger Biopharmaceuticals, received consulting fees from AstraZeneca, is a cofounder with equity in Delphi Diagnostics and issued patents for (1) a method to calculate residual cancer burden and (2) genomic signature to measure sensitivity to endocrine therapy. J.P. reports honoraria from Methods in Clinical Research—faculty SCION workshop; support from ASCO and advocate scholarship; AACR—SSP program; VIVLI, U Wisc SPORE—EAB, QuantumLEAD—COVID DSMB, PCORI—reviewer and I-SPY advocate lead. P.P. reports institutional research funding from Genentech/Roche, Fabre-Kramer, Advanced Cancer Therapeutics, Caris Centers of Excellence, Pfizer, Pieris Pharmaceuticals, Cascadian Therapeutics, Bolt, Byondis, Seagen, Orum Therapeutics and Carisma Therapeutics; consulting fees from Personalized Cancer Therapy, OncoPlex Diagnostics, Immunonet BioSciences, Pfizer, HERON, Puma Biotechnology, Sirtex, CARIS Life sciences, Juniper, Bolt Biotherapeutics and AbbVie; honoraria from DAVA Oncology, OncLive/MJH Life Sciences, Frontiers—publisher, SABCS and ASCO; Speakers' Bureau: Genentech/Roche (past); United States patent no. 8486413, United States patent no. 8501417, United States patent no. 9023362, United States patent no. 9745377; uncompensated roles with Pfizer, Seagen and Jazz. R.A.S. reports institutional research funding from OBI Pharma, QLHC, AstraZeneca and Gilead, serves on AstraZeneca and Stemline Advisory Boards and Gilead Speaker's Bureau and reports consultancy role with QLHC. A.D. reports institutional research funding from Novartis, Pfizer, Genentech and NeoGenomics; Program Chair, Scientific Advisory Committee, ASCO. D.Y. reports research funding from NIH/NCI P30 CA 077598, P01 CA234228-01 and R01CA251600, consulting fees from Martell Diagnostics, and honoraria and travel for speaking at the 'International Breast Cancer Conference.' L.J.v.V. is a founding advisor and shareholder of Exai Bio and is a part-time employee and owns stock in Agendia. N.M.H. reports institutional research funding from NIH. L.J.E. reports funding from Merck and Co., participation on an advisory board for Blue Cross Blue Shield and personal fees from UpToDate and is an unpaid board member of QLHC. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-024-03266-2.

Correspondence and requests for materials should be addressed to Laura J. Esserman.

Peer review information *Nature Medicine* thanks Barbara Pistilli and Sze-Huey Tan for their contribution to the peer review of this work. Primary Handling Editor: Ulrike Harjes, in collaboration with the *Nature Medicine* team.

nature portfolio

Corresponding author(s): Laura Esserman

Last updated by author(s): 2024.08.07

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	OpenClinca 4 (Stack 19.1) electronic data capture system was used at all sites; OneSource integration with clinical sites' EHR systems was deployed at a number of centers to collect laboratory information to electronic CRFs.
Data analysis	Efficacy analysis of Block A alone was performed using R version 4.1.3 and STAN 2.21.0 via the RSTAN package (2.21.5), which implements the NUTS sampler. Efficacy analysis of treatment strategy was performed using R 4.2.2 via the stats package (4.2.2).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

De-identified subject level data and/or clinical specimens are made available to members of the research community upon approval of the I-SPY Data Access and Publications Committee. Details of the application and review process are available at: https://www.quantumleaphealth.org/for-investigators/clinicians-proposal-

submissions/. I-SPY aims to make complete patient-level clinical datasets available for public access within 6 months of publication, as the data is complex and requires significant annotation to ensure its usability.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	The I-SPY2.2 trial in early breast cancer enrolls adult women.
Reporting on race, ethnicity, or other socially relevant groupings	Self-reported race and ethnicity of the study population are reported in the manuscript. These variables are not controlled for in the efficacy analysis.
Population characteristics	The Dato-DXd arm was open to patients diagnosed with primary early stage high-risk HER2-negative breast cancer. The enrolled population included 58.3% White, 8.7% Asian, and 10.7% Black and 11.7% Hispanic/Latino patients. Median age at screening was 48 (range 28-78)
Recruitment	Participants are invited to screen for I-SPY2.2 at the invitation of local clinical personnel at the time of breast cancer diagnosis at open study sites across the US (32 during the arm's tenure). Participants did not receive compensation for trial participation. I-SPY2.2 is a complex trial opened at academic centers that requires frequent (clinical/imaging) visits. The patient population may self-select towards those who are more highly educated or those who are physically able to adhere to the (relatively more demanding) visit schedule. These biases are unlikely to impact comparisons to a fixed threshold for efficacy of Block A alone. The efficacy of the treatment strategy in comparison to the dynamic control built from I-SPY2 data may be more susceptible to these biases. But I-SPY2 prior to the design change is a complex trial in itself requiring frequent imaging/clinical visits, so these biases are also unlikely to materially impact findings.
Ethics oversight	Wake Forest School of Medicine (central IRB), independent DSMB

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was set to 100 patients. This is set via simulation of the trial's operating characteristics for comparing the efficacy of Block A relative to a fixed pCR rate threshold. We used a simple Bayesian logistic model for pCR (as a function of treatment) in these simulations. This is a limitation; but simulations using the covariate adjusted models over a large range of parameter assumptions was not tractable. Via our simulations, we recognize that graduation may be difficult to achieve in the smaller subtypes given the overall 100 patient sample size, particularly in the Immune-DRD+ subtype which is compared against the 40% threshold. However, we note that our sample size is slightly larger than the maximum accrual size for I-SPY 2 for agents opening in HER2- (or HER2+) patients; and was deemed adequate for identifying agents with large efficacy signals (relative to the thresholds).
Data exclusions	None
Replication	n/a (clinical trial). Each patient is unique and can only be treated once and assessed once for trial endpoints. No replication is possible.
Randomization	When the Dato-DXd arm was open to enrollment, participants were randomized with equal probability to open arms at Block A for their subtype. The number of open Block A arms in each subtype varied over the course of Dato-DXd enrollment. In Block B, participants with the Immune-DRD+ subtype were randomized 1:1 to receive one of two Block B treatment courses: paclitaxel plus carboplatin, or paclitaxel plus carboplatin and pembrolizumab. For all other subtypes, Block B treatment was assigned to a single available course of treatment for this block.
Blinding	Investigators are not blinded to randomization results, but are blinded to efficacy analysis findings until announcement that experimental regimens have exited the trial. Blinding to group allocation is not possible because experimental arms may have unique safety profile which requires specific toxicity management strategies and/or dose modifications.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	clinicaltrials.gov identifier NCT01042379
Study protocol	Uploaded with the manuscript
Data collection	103 patients were randomized to this study arm of I-SPY2.2 across 32 (28 main and 16 satellite) study sites between September 2022 and August 2023. CLinical sites were as follows: University of California San Francisco; University of California San Diego; UC Davis Health; Yale University; Columbia University; Oregon Health & Science University; University of Pennsylvania; University of Chicago; University of Minnesota; University of Colorado; University of Alabama Birmingham; Mayo Clinic, Rochester; Hoag Family Cancer Institute; University of Utah (Huntsman Cancer Institute); Cooperman Barnabas Medical Center; Rutgers Cancer Institute of New Jersey; Loyola University Medical Center; Emory University; Georgetown University; City of Hope (Lennar); University of Southern California; Sparrow Hospital; Wake Forest University; Sanford Health
Outcomes	Primary endpoint is pathologic complete response. Primary endpoint is assessed using the Residual Cancer Burden (RCB) method. Distributions of pCR rates after Block A alone are compared to set subtype-specific thresholds. Distribution of pCR rate of treatment strategy across all blocks is compared to probability distributions based on I-SPY2 arms reflecting (current and historical) standard of care. Secondary endpoints are RCB assessed at the time of surgery using the RCB method and event-free survival collected/updated yearly using case report forms (first beginning 30 days after study completion).

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*