新薬と臨牀 2023；72：841-864

ORIGINAL ARTICLE

Real-world Clinical Outcomes in High-risk HR＋/HER2－Early Breast Cancer in Japan

――A Study Using Nationwide Hospital-based Administrative Claims and Electronic Medical Record Databases――

Norihiko Inoue[1], Takuaki Tani[1], Natsuko Kanazawa[1], Eriko Tokunaga[2], Kenjiro Aogi[3], Zhihong Cai[4], Satoshi Osaga[4], Tsutomu Kawaguchi[4] and Yoshinori Tanizawa[4]

1：*National Hospital Organization Headquarters, Tokyo, Japan*
2：*National Hospital Organization Kyushu Cancer Center, Fukuoka, Japan*
3：*National Hospital Organization Shikoku Cancer Center, Matsuyama, Japan*
4：*Japan Drug Development & Medical Affairs, Eli Lilly Japan K.K., Kobe, Japan*

Corresponding author：Yoshinori Tanizawa
Japan Drug Development & Medical Affairs, Eli Lilly Japan K.K.
5-1-28 Isogami-dori, Chuo-ku, Kobe 651-0086, Japan
Tel：＋81-78-242-9624　E-mail：tanizawa_yoshinori@lilly.com

Abstract

Background：Real-world evidence from Japan for clinical outcomes and treatment patterns in patients with hormone receptor-positive（HR＋）and human epidermal growth factor receptor 2-negative（HER2－）early breast cancer（EBC）is limited, especially in the high-risk population. We aimed to provide recent evidence in this population using a nationwide standardized and established dataset owned by National Hospital Organization in Japan.

Methods：Adult females diagnosed with stage Ⅰ-Ⅲ breast cancer who underwent breast resection surgery during April 2015-March 2020 were included. Patients with ≥N2, or N1 and ≥T3 tumor as the record in DPC at surgery were considered 'high-risk' patients. The primary objective was to evaluate disease-free survival（DFS）and factors affecting DFS. The secondary objective was to describe the characteristics and treatment patterns.

Results：Among 8579 included patients, 366 were high risk. Except cancer stage, all other characteristics were similar between high-risk and overall population. Neoadjuvant therapy（45.4% vs. 7.9%）and adjuvant chemotherapy（36.6% vs. 19.4%）use was more common in high-risk versus overall population. Cyclophosphamide-based regimens were the most frequently used adjuvant chemotherapy regardless of risk

profile. Five-year DFS was lower in high-risk versus overall population（80.4% vs. 95.5%）. High-risk was also the greatest factor affecting cancer recurrence/death in the overall population（hazard ratio 5.0, 95% confidence interval 3.7-7.0；p＜0.001）.

Conclusion：This study demonstrated a significantly higher probability of poor prognosis in high-risk patients with HR＋, HER2－EBC. Therefore, further improvement of treatment outcomes in the high-risk population is desirable（**Fig**：Plain language summary of the study）.

Introduction

Breast cancer is the most commonly diagnosed cancer among women[1], with early breast cancer（EBC）accounting for ＞90% of cases[2]. In Japan, hormone receptor-positive（HR＋）and human epidermal growth factor receptor 2-negative（HER2－）is the most common breast cancer subtype（＞70% of cases）[3]. The standard of care for EBC includes surgery along with radiotherapy, adjuvant or neoadjuvant chemotherapy, and endocrine therapy（ET）based on individual factors[4]. Around 10.4% of patients with EBC（irrespective of HR-status）will experience cancer recurrence[5], however, this proportion increases among patients with HR＋, HER2－EBC and certain high-risk features[6]（approximately 30% to 37%）[7][8].

To address this unmet need of high-risk patients with EBC, the monarchE trial, an open-label, PhaseⅢ, randomized, global trial, which enrolled 5637 high-risk HR＋, HER2－patients with EBC, was conducted. Patients were randomly assigned to standard-of-care adjuvant ET with or without abemaciclib for two years. Abemaciclib＋ET was superior to ET alone for the primary outcome（invasive disease-free survival［IDFS］：hazard ratio［HR］0.75, 95% confidence interval［CI］0.60-0.93；p＝0.01）[9]. Based on the monarchE trial, the United States（US）, the European Union, and Japan approved the use of abemaciclib ＋ET for patients with HR＋, HER2－EBC at a high risk of cancer recurrence[10]-[12].

To understand how such clinical trial results will translate into clinical practice, real-world evidence on clinical outcomes for HR＋, HER2－patients with EBC is essential. However, published evidence is limited, especially in the high-risk population treated with the standard-of-care before monarchE approval. Studies from the US have reported three-fold greater risk of recurrence in high-risk patients versus non-high-risk patients[7], and the treatment patterns in patients with EBC[13]. However, there are no large-scale real-world Japanese studies demonstrating clinical outcomes, especially among high-risk patients with HR＋, HER2－EBC, in the context of the overall journey of patients with HR＋, HER2－EBC. In the current study, we aimed to generate recent evidence by describing real-world patient characteristics, treatment patterns, and clinical outcomes among Japanese patients with HR＋, HER2－EBC, with focus on the high-risk patients among them（based on a monarchE-like definition）[9], using a nationwide standardized and established database.

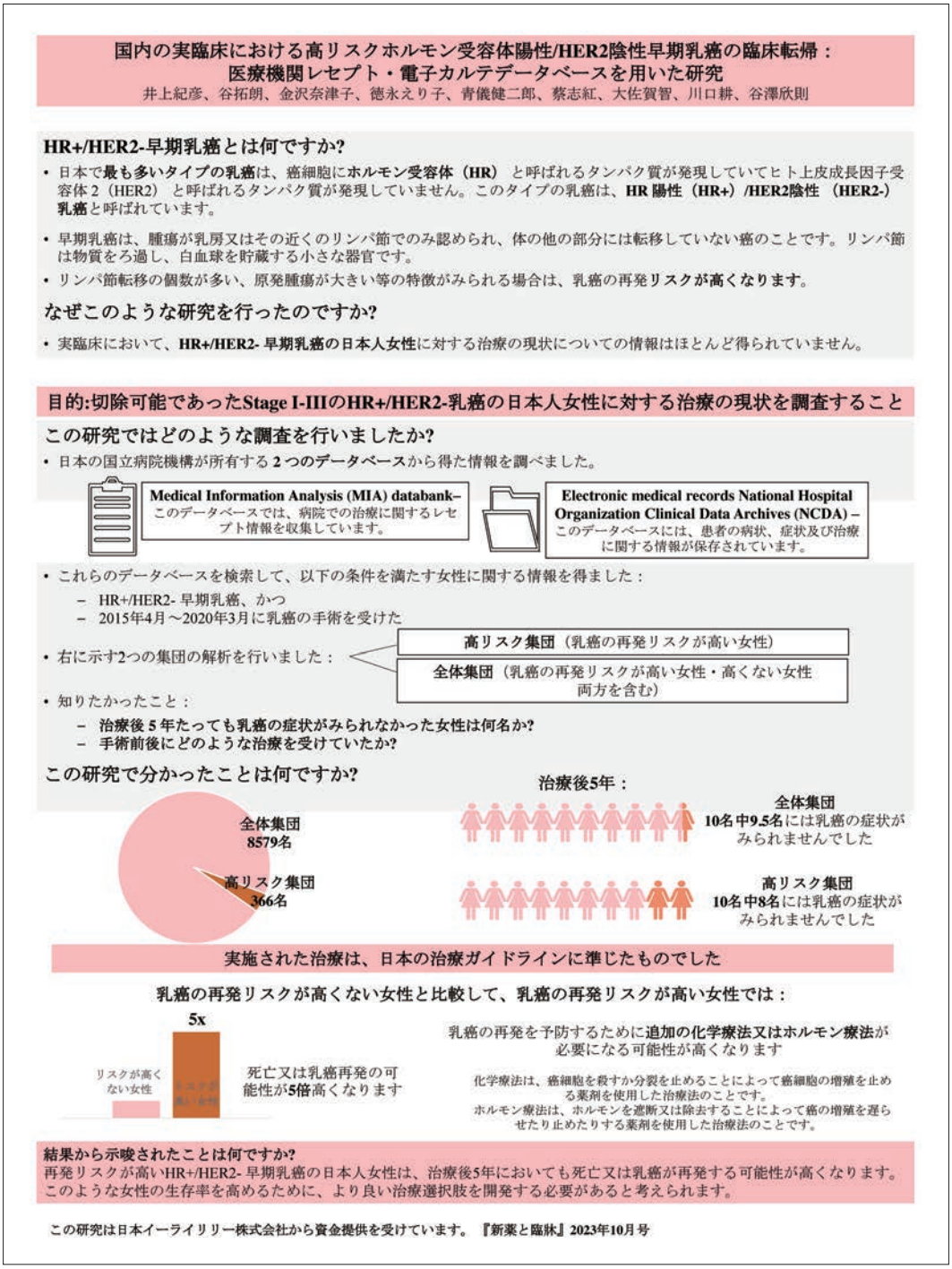


Fig. Plain language summary of the study（Japanese only）

Ⅰ Methods

1. Data source

This retrospective observational study was conducted using data from the hospital-based claims Medical Information Analysis (MIA) databank and the electronic medical record (EMR) National Hospital Organization (NHO) Clinical Data Archives (NCDA) database[14]. Both the databases are owned by the NHO. The MIA database has been collecting administrative claims data from 140 hospitals since April 2010 with more than 8 million inpatients and outpatients[14]. It also includes discharge summaries containing clinical information such as cancer stage, T/N stage, initial versus recurrent flag, and activities of daily living (ADL). The NCDA database includes EMR data regarding the disease, admission/discharge, transfer, outpatient visits, diet, prescriptions, medications, and laboratory test result data for patients since January 2016. Only anonymized data is provided for secondary use, including this study. This study included data from NHO hospitals that participate in both the databases.

2. Study design and patient population

The date of the first breast resection surgery (kubun codes K474-K476) was defined as the index date and this date needed to be in index period (April 1, 2015 to March 31, 2020). The hospitalization including the index date was defined as the index hospitalization. Patients were followed-up for at least 1 year till the last hospital visit up to March 31, 2021.

The study included patients who underwent breast resection surgery in the index period; were diagnosed with stage Ⅰ-Ⅲ (based on tumor, nodes, metastases classification [TNM]) breast cancer (International Classification of Diseases, Tenth Edition [ICD-10] C50, except for sarcoma or HER2 positive cancer)[15] for the index hospitalization; received any endocrine drugs (HR+surrogate); did not receive HER2−targeting drugs (HER2−surrogate); were diagnosed with initial and not recurrent cancer at the index hospitalization; were female and aged ≥18 years at the index date; and had a follow-up duration of ≥1 year or record of death after the index date. Patients who underwent breast resection surgery before the index date; or had a diagnosis of metastasis/recurrence except for axillary lymph node metastasis before the index date were excluded. According to our previous validation study[16], either the clinical or pathological stage retrieved from each hospital matched 83.7% of overall stage, 95.8% of T-classification, and 89.8% of N-classification in the DPC database, thus we utilized the T/N classification retrieved from DPC information as T/N stage at surgery. In this study, ≥N2, or N1 and ≥T3 tumor as the record in DPC data at surgery were defined as 'monarchE-like high-risk criteria', and the subgroup of patients with this criteria as 'high-risk'. This definition is not compatible to but partially match the definition of high-risk criteria used in the monarchE trial[9]. We could not use the exact monarchE criteria since the NHO databases do not record the tumor grade or Ki-67 index.

This observational study only used data collected previously, does not include any intervention, and deidentifies the data to protect patient privacy. Therefore, a formal consent form was not required. This study

was conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki and that are consistent with Good Pharmacoepidemiology Practices and applicable laws and regulations of Japan. This study was approved by the Institutional Review Board of the NHO Headquarters（R3-0614001）.

3. Objectives and variables

We evaluated demographic and clinical characteristics of the patients including age, sex, body mass index, ADL, Charlson Comorbidity Index（CCI）, menopausal status, and tumor stage. All the variables, except menopausal status, were measured during index hospitalization and obtained from the discharge summary. Menopausal status was defined based on the presence of goserelin/leuprorelin prescriptions in the follow-up period as a surrogate. ADL was evaluated using the Barthel Index, a scale to evaluate a person's current level of ability to perform 10 everyday activities[17]. Patients were deemed independent during admission if their Barthel index score was 100 and dependent if they scored＜100.

The primary objective was evaluation of duration of disease-free survival（DFS：defined as time to metastasis or recurrence or death）during the follow-up period in the patients. These events were defined using pre-validated algorithms[16]. Metastasis/recurrence was defined using lasso-based and rule-based algorithms. In the lasso-based algorithm, presence of codes in **Table S1** were used as independent variables for the regression model. The event was defined as present when the prediction score of the lasso regression model was higher than the threshold value that maximized the sum of sensitivity and specificity in the receiver operating characteristics curve, and the event date was defined as the earliest date among the dates with codes in **Table S1**. In the rule-based algorithm, presence of any metastasis/recurrence codes in **Table S1** defined the event, and the event date was defined as the earliest date among those records. Death was defined as the presence of a death record in the discharge summary, EMR, or claims records（**Table S1**）. The earliest date among these events defined the event date.

The secondary objective described the demographic and clinical characteristics and treatment patterns. Use of endocrine and chemotherapeutic drugs over two lines of therapy（LoT）for neoadjuvant and adjuvant therapy were assessed as per the definitions in **Table S1**. Use of post-operative radiation therapy within 180 or 365 days after the index date was assessed too.

4. Statistical analysis

DFS was analyzed using the Kaplan-Meier method with log-rank test and was censored at the last hospital visit for patients who did not have a record of metastasis/recurrence or death. Median（95% CI）of DFS and DFS% at specified time points（1, 2, 3, 4, and 5 years）were estimated. The lasso-based algorithm was considered as the main analysis and rule-based algorithm as the sensitivity analysis. Multivariable Cox regression analysis was performed to assess the risk factors affecting DFS. Values were missing in 0.5% and 1.2% of the BMI and ADL data, respectively. Hence, the missing values were imputed with Multiple Imputation by Chained Equations before the Cox regression analysis[18]. The demographic and clinical characteristics

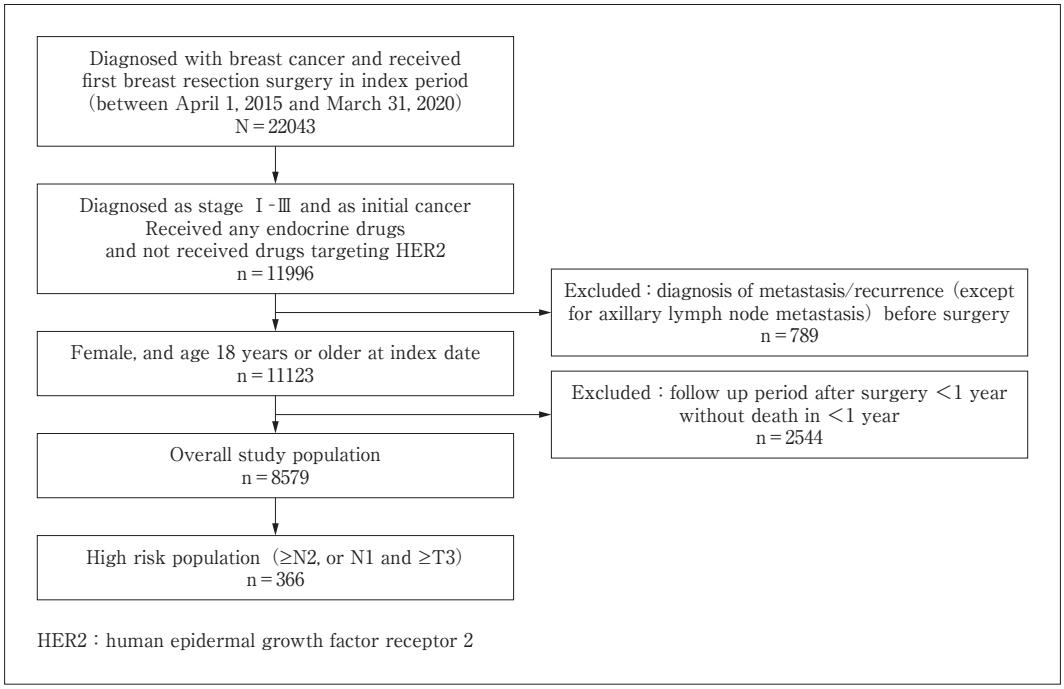


Fig. 1 Selection of the study population

were described descriptively. Treatment patterns were presented using Sankey diagrams.

Similar analyses were performed for 'high-risk' patients for all the outcomes. Subgroup analysis for DFS was performed for patients stratified by stage (Ⅰ vs. Ⅱ vs. Ⅲ), by T-stage (T1 vs. T2 vs. T3 vs. T4), and by N-stage (N0 vs. N1 vs. N2 vs. N3). All statistical analyses were conducted using R 4.1.3 (R core team)[19].

Ⅱ Results

1. Demographic and clinical characteristics

A total of 8579 patients met all the inclusion and exclusion criteria while 366 patients (4.3%) were considered 'high-risk' among them (**Fig. 1**). The median (range) age of the overall study population was 62.0 (24.0-99.0) years and the mean (standard deviation [S.D.]) follow-up period was 3.2 (1.4) years. The majority of the patients were independent (94.8%) in ADL. The most frequent tumor stage was stage Ⅰ (55.6%), followed by stage ⅡA (28.4%), and stage ⅡB (8.3%). High-risk patients reported similar characteristics as the overall population, except for tumor stage. In high-risk patients, stage ⅢB was the most frequent tumor stage (41.5%), followed by stage ⅢA (38.8%), and stage ⅢC (19.7%) (**Table 1**).

2. Peri-operative treatment patterns

1) Neoadjuvant therapies

Neoadjuvant therapy use was more common in high-risk patients than the overall population (1st neoadjuvant : 45.4% vs. 7.9%, 2nd neoadjuvant : 31.4% vs. 4.0%) (**Table 2**). The most common 1st neoadjuvant in high-risk patients was 5-fluorouracil + epirubicin

Table 1 Patient characteristics

Parameter	Overall patients（N＝8579）	High-risk patients（n＝366）
Age in years, median（range）	62.0（24.0-99.0）	63.0（29.0-93.0）
Follow-up period in years, mean（S.D.）	3.2（ 1.4）	3.2（ 1.4）
BMI, mean（S.D.）	23.5（ 4.2）	23.7（ 4.3）
BMI category, n（%）[a]		
＜25.0	5948（69.3）	240（65.6）
≥25.0	2584（30.1）	120（32.8）
Activities of daily living independence, n（%）[b]		
Dependent	348（ 4.1）	29（ 7.9）
Independent	8130（94.8）	329（89.9）
CCI, mean（S.D.）	2.4（ 1.0）	2.7（ 1.2）
CCI＜2, n（%）	6698（78.1）	249（68.0）
Pre-menopausal, n（%）[c]	928（10.8）	38（10.4）
Tumor stage in the database, n（%）		
Ⅰ	4770（55.6）	0（ 0.0）
ⅡA	2437（28.4）	0（ 0.0）
ⅡB	712（ 8.3）	0（ 0.0）
ⅢA	142（ 1.7）	142（38.8）
ⅢB	282（ 3.3）	152（41.5）
ⅢC	72（ 0.8）	72（19.7）
Unknown	164（ 1.9）	0（ 0.0）
Index year, n（%）		
2015	1203（14.0）	47（12.8）
2016	1833（21.4）	88（24.0）
2017	1874（21.8）	81（22.1）
2018	1645（19.2）	77（21.0）
2019	1766（20.6）	64（17.5）
2020	258（ 3.0）	9（ 2.5）

a：Patients with unavailable data, n（%）；Overall patients＝47（0.5%）and high-risk patients＝6（1.6%）.
b：Patients with unavailable data, n（%）；Overall patients＝101（1.2%）and high-risk patients＝8（2.2%）.
c：Menopausal status was defined based on the presence of goserelin/leuprorelin prescriptions in the follow-up period.
N denotes the total study population size and n denotes the number of patients for individual cohorts or parameters.

BMI：body mass index, CCI：charlson comorbidity index, S.D.：standard deviation

Table 2 Neoadjuvant therapies in patients with HR＋, HER2－EBC

	Overall patients（N＝8579）		High-risk patients（n＝366）	
	1st neoadjuvant	2nd neoadjuvant	1st neoadjuvant	2nd neoadjuvant
Patients who received each LoT, n（%）	677（ 7.9）	343（ 4.0）	166（45.4）	115（31.4）
Duration in days, mean（S.D.）	57.9（37.3）	62.1（18.5）	56.9（33.7）	63.4（16.2）
1st common regimen, n	FEC, 141	DTX, 169	FEC, 48	DTX, 60
2nd common regimen, n	LET, 114	FEC, 87	EC, 34	PTX, 27
3rd common regimen, n	EC, 87	PTX, 50	AC, 17	FEC, 20
4th common regimen, n	ANA, 68	LET, 9	LET, 13	LET, 4
5th common regimen, n	nab-PTX, 65	nab-PTX, 8	nab-PTX, 11	BEV＋PTX＋TAM, 1

N denotes the total study population size and n denotes the number of patients for individual cohorts or parameters.

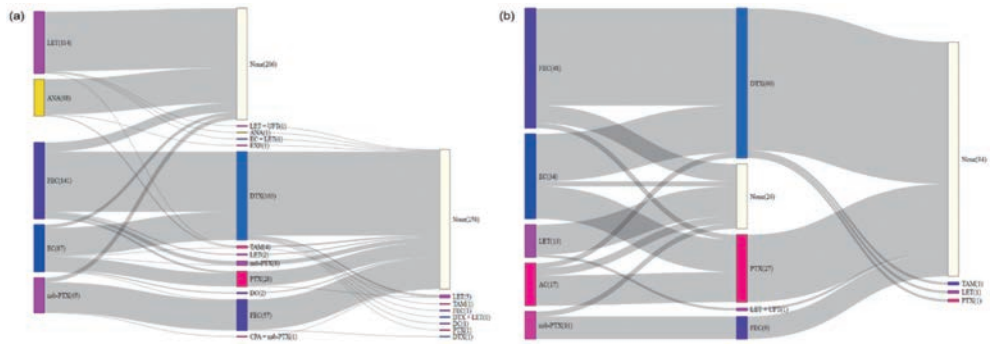
AC：cyclophosphamide＋doxorubicin, ANA：anastrozole, BEV：bevacizumab, DTX：docetaxel, EBC：early breast cancer, EC：epirubicin＋cyclophosphamide, FEC：5-fluorouracil＋epirubicin＋cyclophosphamide, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive, LET：letrozole, LoT：line of therapy, nab-PTX：nanoparticle albumin-bound paclitaxel, PTX：paclitaxel, S.D.：standard deviation, TAM：tamoxifen

＋cyclophosphamide（FEC）followed by epirubicin＋cyclophosphamide（EC）. The most common 1st neoadjuvant in the overall population was FEC followed by letrozole（**Table 2**）.

Apart from reporting the neoadjuvant therapy regimens, we presented the treatment sequence of neoadjuvant therapies in the overall and high-risk populations, starting from the five most common 1st neoadjuvant regimens, using a Sankey plot（**Fig. 2**a and **2**b）. In both the populations, most patients received up to two neoadjuvant therapies；only a few patients received a 3rd neoadjuvant therapy.

2）Adjuvant therapies in the overall study population

Among the overall study population（n＝8579）, 6445 patients（75.1%）initiated any adjuvant therapy within 60 days after surgery. Of those, 1253 patients（19.4%）first received chemotherapy as adjuvant therapy. The most common 1st adjuvant chemotherapy was cyclophosphamide＋docetaxel（DC）（n＝536）followed by EC（n＝243）（**Table 3**）. Among the 1253 patients who received adjuvant chemotherapy, 860（68.6%）further received adjuvant ET within 60 days after adjuvant chemotherapy. The most common 1st adjuvant ET after adjuvant chemotherapy was tamoxifen（n＝373）followed by



Treatment sequence including the five most common 1st neoadjuvant therapy and subsequent lines of neoadjuvant therapy in the overall population (**Fig. 2**a) and high-risk population (**Fig. 2**b).

The numbers of patients for each regimen in the subsequent lines of therapy differs from those in **Table 2** because only the treatment sequences starting from the five most common 1st neoadjuvant regimens are shown here.

5-FU：5-fluorouracil, ANA：anastrozole, CPA：cyclophosphamide, DOX：doxorubicin, DTX：docetaxel, EBC：early breast cancer, EPI：epirubicin, EXE：exemestane, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive, LET：letrozole, nab-PTX：nanoparticle albumin-bound paclitaxel, PTX：paclitaxel, TAM：tamoxifen, UFT：uracil-tegafur

Fig. 2　Neoadjuvant treatment sequence in patients with HR＋, HER2－EBC

anastrozole（n＝254）（**Table 3**）.

Among the 6445 patients who received any adjuvant therapy within 60 days after surgery, 5192 patients（80.6%）started adjuvant ET without chemotherapy（**Table 4**）. The most common 1st adjuvant ET in these patients was anastrozole（n＝2038）followed by tamoxifen（n＝1744）.

The treatment sequence of adjuvant therapies in the overall population were presented using Sankey diagrams（**Figs. 3**a, **3**b, and **4**a）. From the five most common 1st adjuvant chemotherapy（n＝1127）, some patients continued onto the 2nd adjuvant chemotherapy. The most common chemotherapy sequence was from EC or FEC to docetaxel（DTX）or paclitaxel（PTX）（**Fig. 3**a）. Major subsequent adjuvant ETs after the major adjuvant chemotherapies were tamoxifen（31.5%；355/1127）, anastrozole（20.7%；233/1127）, and letrozole（16.9%；191/1127）（**Fig. 3**b）. Among patients who started the five most common ETs without chemotherapy as adjuvant therapy, 12.1%（626/5157）received the second adjuvant ET（**Fig. 4**a）.

3）Adjuvant therapies in the high-risk population

Among the high-risk patients（n＝366）, 284 patients（77.6%）initiated any adjuvant therapy within 60 days after surgery. Of those, 104 patients（36.6%）first received chemotherapy as the adjuvant therapy（**Table 3**）. The most common 1st adjuvant chemotherapy was FEC（n＝32）followed by EC（n＝21）（**Table 3**）. Among the 104 patients who received adjuvant chemotherapy, 72（69.2%）further received adjuvant ET within 60 days after adjuvant chemotherapy（**Table 3**）. The most common 1st adjuvant

Table 3 Adjuvant therapies in patients with HR＋, HER2－EBC who received adjuvant chemotherapy

	Overall patients with adjuvant chemotherapy（n＝1253）				High-risk patients with adjuvant chemotherapy（n＝104）			
	1st adjuvant CT	2nd adjuvant CT	1st adjuvant ET	2nd adjuvant ET	1st adjuvant CT	2nd adjuvant CT	1st adjuvant ET	2nd adjuvant ET
Patients who received each LoT, n（%）	1253 (100)	514 (41.0)	860 (68.6)	66 (5.3)	104 (100)	65 (62.5)	72 (69.2)	5 (4.8)
Duration in days, mean（S.D.）	97.8 (197.3)	96.7 (183.0)	953.3 (555.1)	496.9 (465.1)	100.3 (192.1)	76.4 (99.7)	872.9 (512.8)	272.0 (548.5)
1st common regimen, n	DC, 536	DTX, 318	TAM, 373	ANA, 19	FEC, 32	DTX, 42	ANA, 23	ANA, 2
2nd common regimen, n	EC, 243	PTX, 122	ANA, 254	TAM, 11	EC, 21	PTX, 15	LET, 23	ABE＋FUL, 1
3rd common regimen, n	FEC, 235	FEC, 31	LET, 206	LET, 10	AC, 16	FEC, 2	TAM, 23	EXE, 1
4th common regimen, n	AC, 82	EC, 8	EXE, 9	EXE, 9	DC, 14	DC, 1	EXE, 2	FUL, 1
5th common regimen, n	DTX, 31	DC, 4	TOR, 9	TOR, 7	DTX, 6	EC, 1	TOR, 1	－

ABE：abemaciclib, AC：cyclophosphamide＋doxorubicin, ANA：anastrozole, CT：chemotherapy, DC：cyclophosphamide＋docetaxel, DTX：docetaxel, EBC：early breast cancer, EC：epirubicin＋cyclophosphamide, ET：endocrine therapy, EXE：exemestane, FEC：5-fluorouracil＋epirubicin＋cyclophosphamide, FUL：fulvestrant, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive, LET：letrozole, LoT：line of therapy, PTX：paclitaxel, S.D.：standard deviation, TAM：tamoxifen, TOR：toremifene

ET after adjuvant chemotherapy was anastrozole, letrozole, and tamoxifen（all n＝23）. The mean（S.D.）duration of 1st adjuvant ET was shorter in high-risk patients compared to the overall study population（872.9［512.8］days vs. 953.3［555.1］days）.

Among the 284 high-risk patients who received any adjuvant therapy within 60 days after surgery, 180 patients（63.4%）received adjuvant ET without chemotherapy（**Table 4**）. The most common 1st adjuvant ET in the high-risk population was letrozole（n＝65）followed by anastrozole（n＝61）. The mean（S.D.）duration of 1st adjuvant ET was shorter in high-risk patients compared to the overall study population（866.0［581.8］days vs. 980.7［584.3］days）.

The treatment sequences of adjuvant therapies in the high-risk population were presented using Sankey plots. Starting from the five most common 1st adjuvant chemotherapy（n＝89）, a greater proportion of these high-risk patients received the 2nd adjuvant chemotherapy compared to the overall study population. Similar to the overall population, the predominant chemotherapy

Table 4 Adjuvant endocrine therapies in patients with HR+, HER2−EBC who started adjuvant ET therapy without chemotherapy

	Overall patients who started adjuvant ET therapy without chemotherapy (n=5192)		High-risk patients who started adjuvant ET therapy without chemotherapy (n=180)	
	1st adjuvant ET	2nd adjuvant ET	1st adjuvant ET	2nd adjuvant ET
Patients who received each LoT, n (%)	5192 (100)	626 (12.1)	180 (100)	40 (22.2)
Duration in days, mean (S.D.)	980.7 (584.3)	532.4 (497.5)	866.0 (581.8)	421.3 (445.5)
1st common regimen, n	ANA, 2038	TAM, 189	LET, 65	LET, 9
2nd common regimen, n	TAM, 1744	ANA, 135	ANA, 61	ANA, 8
3rd common regimen, n	LET, 1231	EXE, 93	TAM, 45	FUL, 7
4th common regimen, n	EXE, 106	LET, 92	EXE, 4	TAM, 6
5th common regimen, n	TOR, 38	TOR, 40	TOR, 3	EXE, 5

ANA：anastrozole, EBC：early breast cancer, EXE：exemestane, FUL：fulvestrant, HER2−：human epidermal growth factor receptor 2 negative, HR+：hormone receptor-positive, LET：letrozole, LoT：line of therapy, S.D.：standard deviation, TAM：tamoxifen, TOR：toremifene

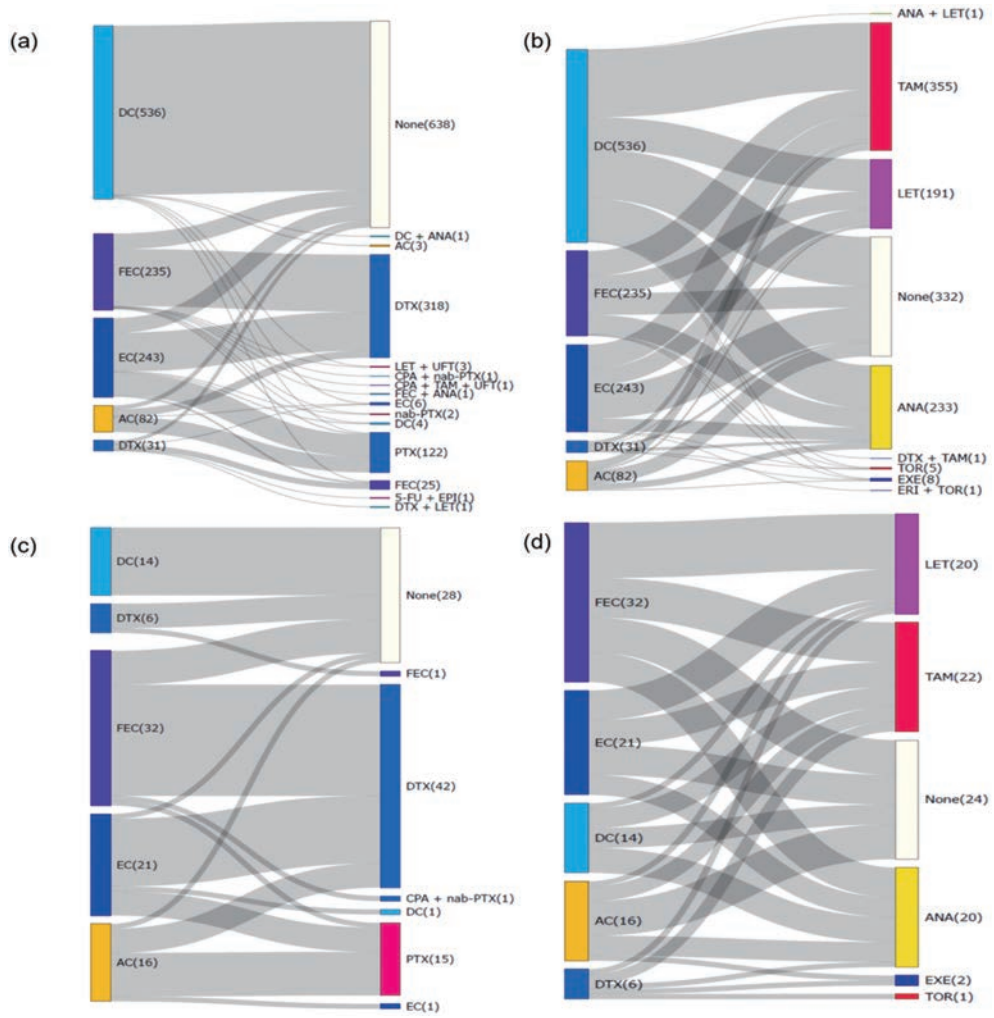
sequence in the high-risk population was from EC or FEC to DTX or PTX as 2nd adjuvant chemotherapy (**Fig. 3**c). The major ETs after the major adjuvant chemotherapies were tamoxifen (24.7%；22/89) and anastrozole and letrozole (for both 22.5%；20/89) (**Fig. 3**d). Among the patients who started the five most common ET without chemotherapy as adjuvant therapy, 21.9% (n=39/178) received the 2nd adjuvant endocrine therapy (**Fig. 4**b).

4) Post-operative radiotherapy

Post-operative radiotherapy use was similar within 180 days for high-risk patients (43.4%) and the overall population (41.9%). However, evaluation for radiotherapy use within 365 days showed higher usage by high-risk patients than the overall population (57.7% vs. 47.0%). No total body irradiation, electromagnetic thermal therapy, or interstitial irradiation were recorded within the evaluated periods.

3. Disease-free survival

Fig. 5 shows the lasso-based algorithm's DFS for the overall study population and high-risk patients. The DFS% was lower for high-risk patients versus the overall study population for five years, ranging from 95.3% vs. 99.3% at 1-year to 80.4% vs. 95.5% at 5-years. Cox regression analysis showed the risk of cancer recurrence was highest in patients with high-risk (HR 5.0, 95% CI 3.7-



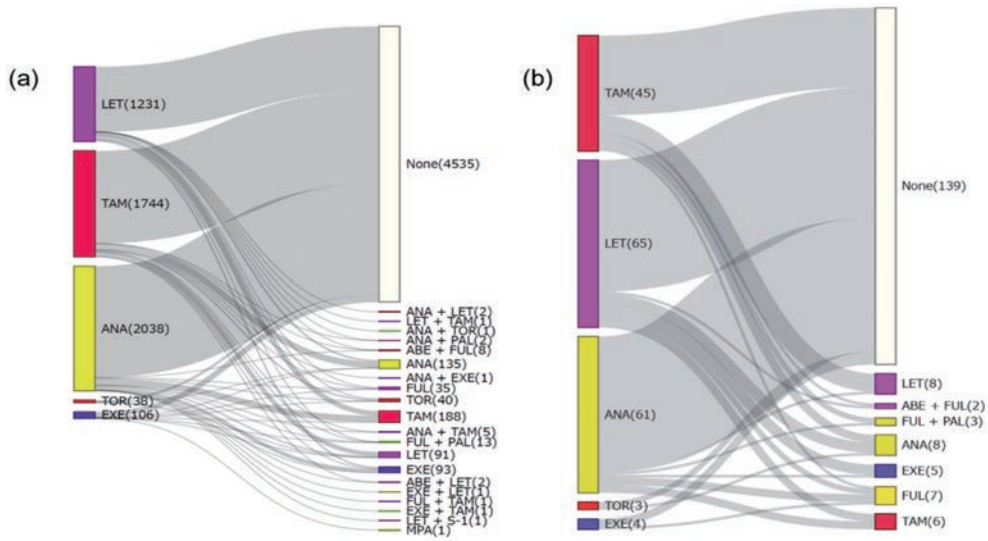
Treatment sequences including the five most common 1st adjuvant chemotherapies and subsequent lines of therapy among patients who received adjuvant chemotherapy in the overall population (**Fig. 3**a and **3**b) and high-risk population (**Fig. 3**c and **3**d).

Among the subsequent lines of therapy, **Figs. 3**a and **3**c only show the 2nd chemotherapy regimens that started within 60 days after the end of the 1st chemotherapy, and **Figs. 3**b and **3**d only show the endocrine therapy regimens that started within 60 days after the end of chemotherapy.

The numbers of patients for each regimen in the subsequent line of therapy differs from those in **Table 3** because only the treatment sequences starting from the five most common 1st neoadjuvant endocrine regimens are shown here.

5-FU：5-fluorouracil, ABE：abemaciclib, ANA：anastrozole, CBDCA：carboplatin, CPA：cyclophosphamide, DOX：doxorubicin, DTX：docetaxel, EBC：early breast cancer, EPI：epirubicin, EXE：exemestane, FUL：fulvestrant, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive, LET：letrozole, nab-PTX：nanoparticle albumin-bound paclitaxel, PAL：palbociclib, PTX：paclitaxel, TAM：tamoxifen, TOR：toremifene, UFT：uracil-tegafur

Fig. 3 Adjuvant treatment sequence in patients with HR＋, HER2－EBC who received adjuvant chemotherapy



Treatment sequences including the five most common 1st adjuvant endocrine therapies and subsequent line of endocrine therapy among patients who did not receive any adjuvant chemotherapy, in the overall population (**Fig. 4**a) and high-risk population (**Fig. 4**b).

The number of patients for each regimen in the subsequent line of therapy differs from those in **Table 4** because only the treatment sequences starting from the five most common 1st neoadjuvant endocrine regimens are shown here.

ABE：abemaciclib, ANA：anastrozole, EBC：early breast cancer, EXE：exemestane, FUL：fulvestrant, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive, LET：letrozole, MPA：medroxyprogesterone, PAL：palbociclib, S-1：tegafur/gimeracil/oteracil, TAM：tamoxifen, TOR：toremifene

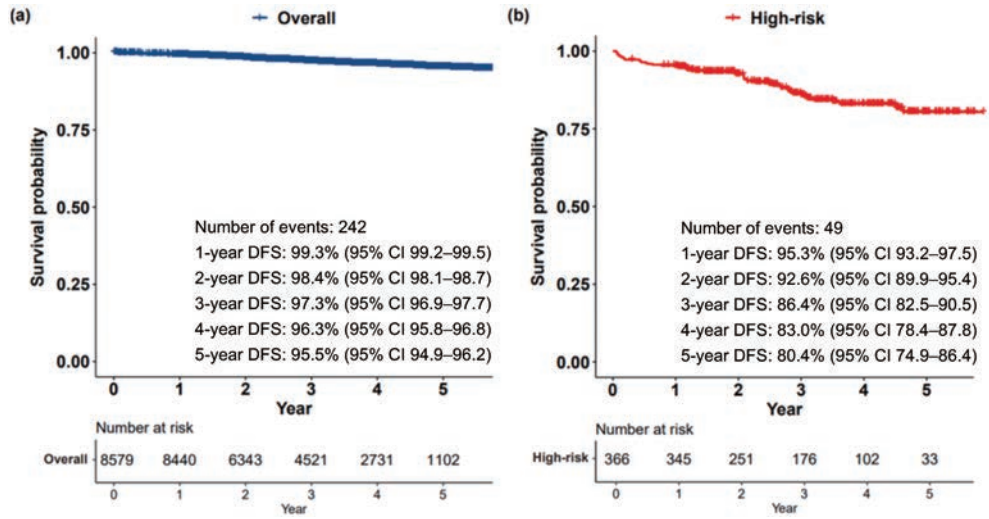
Fig. 4 Adjuvant endocrine therapies in patients with HR＋, HER2－EBC who did not receive adjuvant chemotherapy

7.0), followed by those with CCI ≥3 (HR 1.8, 95% CI 1.4-2.3), and those who received neoadjuvant chemotherapy (HR 1.8, 95% CI 1.3-2.4) (all $p<0.001$; **Fig. 6**). **Fig. S1**a shows the DFS% from subgroup analysis based on the cancer stage, T-stage, and N-stage in the overall population, which showed poorer prognosis in patients with higher stages (**Fig. S1**b-**S1**d).

Similar results were observed using the rule-based algorithm, although the observed DFS% was numerically different from the lasso-based model. The DFS% with the rule-based algorithm was lower for high-risk patients versus the overall study population for five years (**Fig. S2**), and cox regression analysis and subgroup analysis results were also consistent with observations from analysis with the lasso-based algorithm (data not shown).

Ⅲ Discussion

Enhanced knowledge about recurrence risks in real-world patients with HR＋, HER2－EBC can enable physicians to identify and treat patients who require additional therapy and avoid overtreatment in those with low



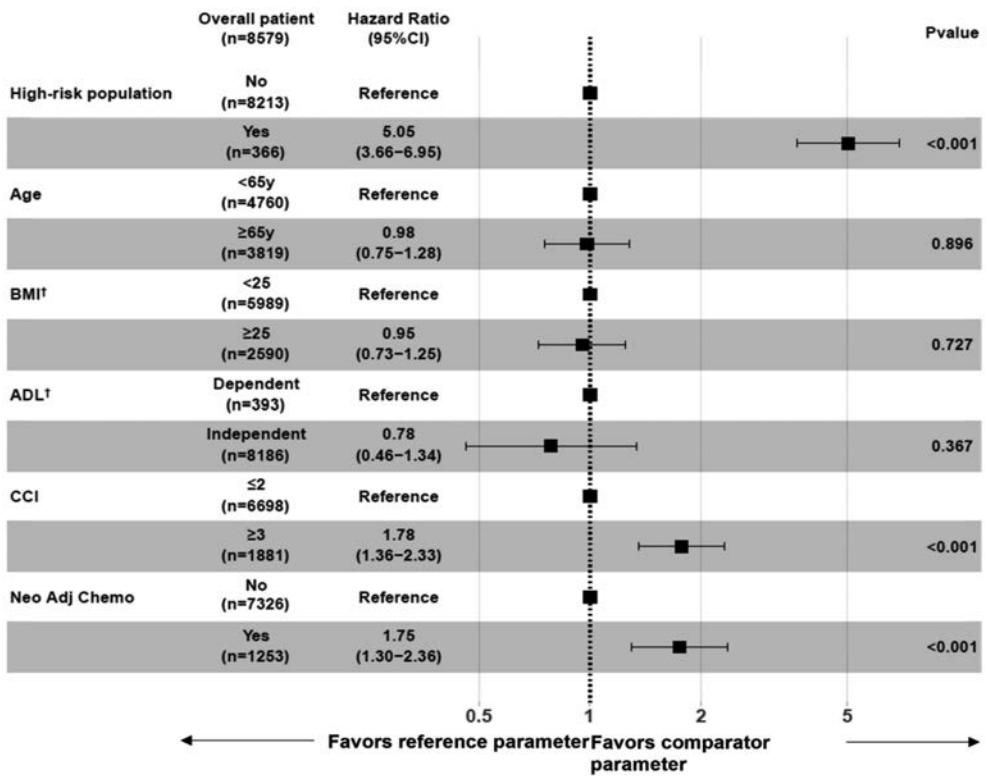
Disease-free survival was analyzed using the Kaplan-Meier method with Log-rank test.
Figs. 5a and **5**b represent lasso-based disease-free survival of the overall and high-risk patients, respectively.

CI：confidence interval, DFS：disease-free survival, EBC：early breast cancer, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive

Fig. 5 Disease-free survival of overall patients and high-risk patients with HR＋, HER2－EBC（lasso-based）

risk of recurrence[7)20)]. This study provides real-world information on DFS and treatment patterns among Japanese patients with HR＋, HER2－EBC, with focus on the high-risk population. Among the overall population, 4.3% met the clinicopathological criteria for high risk of recurrence（≥N2, or N1 and ≥ T3）. As described in the **Study design and patient population** part of **Methods** section, this high-risk definition partially matches the definition used for the monarchE eligibility criteria, but due to the nature of the current study's data source（i.e., absence of histological grade and Ki-67 index）, we could not completely replicate（mimic）the monarchE eligibility criteria[9)]. The high-risk patients had worse 5-years DFS than the overall population（80.4% vs. 95.5%）.

Regarding peri-operative treatment patterns, adjuvant chemotherapy use（including anthracycline- or taxane-based regimens）was more common in high-risk patients than the overall population for both 1st adjuvant chemotherapy（36.6% vs. 19.4% among those with adjuvant therapy）and 2nd adjuvant chemotherapy（62.5% vs. 41.0% among those who received 1st adjuvant chemotherapy）. Among the patients who received adjuvant chemotherapy, the proportion of the patients who further received adjuvant ET within 60 days after chemotherapy was similar between high-risk patients and the overall population（69.2% vs. 68.6%）. Interestingly, the duration of ET was slightly shorter in the high-risk population than in the overall population. The shorter duration could be attributed to the definition of ET end date, which mentions that the end



Hazard ratios and 95% confidence intervals for each factor are presented.
†: Missing values in the ADL and BMI were imputed with Multiple Imputation by Chained Equations and then multivariable Cox regression was performed.

ADL：activities of daily living, BMI：body mass index, CCI：charlson comorbidity index, CI：confidence interval, EBC：early breast cancer, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive, Neo Adj Chemo：neoadjuvant chemotherapy

Fig. 6 Risk factors affecting disease-free survival of overall patients with HR＋, HER2－EBC（lasso-based）

date could be "one day before any other breast cancer drug that was not a part of the regimen was started". Overall, the peri-operative treatment patterns observed were consistent with the Japanese Breast Cancer Society's clinical practice guidelines[4].

Real-world studies from the US have reported IDFS（defined as time from the initiation of ET to the date of earliest event [locoregional recurrence, distant recurrence or death])[7] or 5-year mortality rate[20], as per the monarchE clinicopathological inclusion criteria. In these studies, the 5-year IDFS rate and 5-year mortality rate for patients meeting monarchE criteria versus not meeting the criteria was 70.2% vs. 90.9% and 16.5% vs. 7.0%, respectively. This is consistent with the current study, wherein high-risk patients with HR＋, HER2－EBC had lower 5-year DFS than the overall population（80.4% vs. 95.5%）. The overall population's 5-year DFS in our study（95.5%）was also similar to a

prior Japanese study（recurrence-free survival 94.5%）[21]. We also evaluated DFS based on tumor, node, and cancer stages using both lasso-based and rule-based algorithms. As expected, DFS was poorer in patients with severe disease, thus supporting the clinical validity of NHO databases used in the current study.

Using multivariable Cox regression analysis, we identified the high-risk criteria used in the current study as a strong independent factor in predicting cancer recurrence. Other factors included CCI ≥3 and receiving neoadjuvant chemotherapy. Thus, monarchE-like clinicopathological high-risk criteria can be useful to identify patients with a poor prognosis during real-world clinical practice too. Other real-world studies have identified older age, Black race, pre-menopausal status, Eastern Cooperative Oncology Group Performance Status score 1 or ≥2, Oncotype DX Breast Recurrence Score＞25, tumor size, involved nodal status, histological grades, and history of hormone therapy as prognostic factors for recurrence[7][8][20]-[22]. Our real-world study has contributed to further confirmation of such criteria for identification of high-risk patients. Employing these criteria in clinical practice may facilitate identifying patients who require therapeutic advancements to improve their clinical outcomes.

Limitations

Use of the nation-wide, large-scale NHO databases is a major strength of this study. These databases cover＞8 million patients and are one of the well-established administrative claims/EMR databases in Japan. They have also been used across multiple therapeutic domains, thus showing their usefulness[23]-[25]. Additionally, the outcome definitions and algorithms used in this study have been validated earlier[16], thus demonstrating its reliability. However, a few limitations should be considered while interpreting the results of this study. Metastasis/recurrence events were defined using claims and EMR database records, that were originally created for administrative purposes, therefore they may not fully reflect the true clinical status of the patients. These limitations were mitigated by conducting a validation study before the current study, to understand the algorithm's sensitivity and positive prediction value[16]. Some clinical variables such as menopausal status, HR status, and HER2 status were not included in the databases. Therefore, they were defined by surrogate variables such luteinizing hormone-releasing hormone agonists, ET, and drugs targeting HER2, respectively. Another limitation is how to define the "high-risk" patients with EBC. Definition of high-risk EBC varies in each clinical trial. In this study, a subgroup analysis was conducted with the high-risk criteria for monarchE in mind. However, since it was difficult to identify a high-risk EBC population in the NHO database that perfectly matched those of monarchE due to the limitation of data source, the T/N stage of the DPC was used for convenience. As described in the **Study design and patient population** part of the **Methods** section, the T/N stage of the DPC data was generally consistent with the clinical or pathologic stage at the time of surgery as documented in the EMR. Since both MIA and NCDA databases are owned by the NHO, only hospitals affiliated with the NHO i.e. large,

relatively well-resourced hospitals, participated in this study. Hence, generalizability of the results to other hospitals needs further evaluation. Finally, as the MIA and NCDA are hospital-based databases, patients only be followed within the same hospital, and were lost to follow-up if they moved to other hospitals in or outside the NHO network. Care provided outside the NHO network, if any, could not be detected in these databases.

Conclusion

This study shows that high-risk patients with HR＋, HER2－EBC have a higher probability of recurrence/death than non-high-risk patients. Moreover, the monarchE-like clinicopathological criteria was shown to be useful in identifying such high-risk patients in real-world clinical practice. The current treatment patterns in patients with HR＋, HER2－EBC are consistent with Japanese guidelines, however, therapeutic advancements are needed to provide better clinical outcomes, especially for high-risk patients.

Acknowledgements

This study and its publication were funded by Eli Lilly Japan K.K. Leo J. Philip Tharappel of Eli Lilly Services India Private Limited, Bengaluru, India provided medical writing and editorial support, which was funded by Eli Lilly Japan K.K. All named authors meet the International Committee of Medical Journal Editors criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Disclosures

Norihiko Inoue, Takuaki Tani, and Natsuko Kanazawa are employees of the National Hospital Organization, which received joint research fees from Eli Lilly and Company and report no other competing interest in this work. Eriko Tokunaga received personal fees as honoraria from Chugai Pharmaceutical, AstraZeneca, Daiichi Sankyo, Eli Lilly Japan. Kenjiro Aogi received personal fees as honoraria from Chugai Pharmaceutical, Eisai, AstraZeneca, Taiho Pharmaceutical, Novartis Pharma, Daiichi Sankyo, Mochida Pharmaceutical, Ono Pharmaceutical, Pfizer and Eli Lilly Japan, and his institution received research funds from Chugai Pharmaceutical, Eisai, and Takeda Pharmaceutical. Zhihong Cai, Satoshi Osaga, Tsutomu Kawaguchi, and Yoshinori Tanizawa are employees of Eli Lilly Japan K.K. and hold stock and/or stock options of Eli Lilly and Company.

References

1) Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020：GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021；**71**：209-49.
2) Waks AG, Winer EP. Breast Cancer Treatment：A Review. *JAMA*. 2019；**321**：288-300.
3) Ushimado K, Kobayashi N, Hikichi M, et al. Differences in clinicopathologic features and subtype distribution of invasive breast cancer between women older and younger than 40 years. *Fujita Med J*. 2019；**5**：92-7.
4) Shimoi T, Nagai SE, Yoshinami T, et al. The Japanese Breast Cancer Society Clinical Practice Guidelines for systemic treatment of breast cancer, 2018 edition. *Breast Cancer*. 2020；**27**：322-31.

5) Colleoni M, Sun Z, Price KN, et al. Annual Hazard Rates of Recurrence for Breast Cancer During 24 Years of Follow-Up：Results From the International Breast Cancer Study Group Trials I to V. *J Clin Oncol*. 2016；**34**：927-35.
6) Garutti M, Griguolo G, Botticelli A, et al. Definition of High-Risk Early Hormone-Positive HER2-Negative Breast Cancer：A Consensus Review. *Cancers*（*Basel*）. 2022；**14**：1898.
7) Sheffield KM, Peachey JR, Method M, et al. A real-world US study of recurrence risks using combined clinicopathological features in HR-positive, HER2-negative early breast cancer. *Future Oncol*. 2022；**18**：2667-82.
8) Cheng L, Swartz MD, Zhao H, et al. Hazard of recurrence among women after primary breast cancer treatment--a 10-year follow-up using data from SEER-Medicare. *Cancer Epidemiol Biomarkers Prev*. 2012；**21**：800-9.
9) Johnston SRD, Harbeck N, Hegg R, et al. Abemaciclib Combined With Endocrine Therapy for the Adjuvant Treatment of HR+, HER2−, Node-Positive, High-Risk, Early Breast Cancer（monarchE）. *J Clin Oncol*. 2020；**38**：3987-98.
10) Giordano SH, Freedman RA, Somerfield MR, et al. Abemaciclib With Endocrine Therapy in the Treatment of High-Risk Early Breast Cancer：ASCO Optimal Adjuvant Chemotherapy and Targeted Therapy Guideline Rapid Recommendation Update. *J Clin Oncol*. 2022；**40**：307-9.
11) Verzenio. European Medicines Agency. 2022［updated April 21, 2022］.
Available from：https://www.ema.europa.eu/en/medicines/human/EPAR/verzenios
12) New Drugs Approved in FY 2021：Pharmaceuticals and Medical Devices Agency；2021.
Available from：https://www.pmda.go.jp/files/000250462.pdf
13) Brown J, Scardo S, Method M, et al. A real-world retrospective study of the use of Ki-67 testing and treatment patterns in patients with HR+, HER2− early breast cancer in the United States. *BMC Cancer*. 2022；**22**：502.
14) Kanazawa N, Tani T, Imai S, et al. Existing Data Sources for Clinical Epidemiology：Database of the National Hospital Organization in Japan. *Clin Epidemiol*. 2022；**14**：689-98.
15) Japanese Breast Cancer Society. *General rules for clinical and pathological recording of breast cancer*（*in Japanese*）. *18th ed*. Tokyo：Kanehara Shuppan；2018.
16) Inoue N, Tani T, Kanazawa N, et al. SA71 Validation Study of Algorithms to Identify Cancer Metastasis/Recurrence in Patients With High-Risk HR+/HER2− Early Breast Cancer（EBC）Using Japan Nationwide Hospital-Based Databases Owned By National Hospital Organization. *Value in Health*. 2022；**25**：S497.
17) Wade DT, Collin C. The Barthel ADL Index：a standard measure of physical disability？. *Int Disabil Stud*. 1988；**10**：64-7.
18) van Buuren S, Groothuis-Oudshoorn K. mice：Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*. 2011；**45**：1-67.
19) R Core Team. R：A language and environment for statistical computing. 4.2.0 ed. Vienna, Austria：R Foundation for Statistical Computing；2022.
20) Nelson DR, Brown J, Morikawa A, Method M. Breast cancer-specific mortality in early breast cancer as defined by high-risk clinical and pathologic characteristics. *PLoS One*. 2022；**17**：e0264637.
21) Yamada Y, Mukai H, Tokudome Y, et al.

Improved overall survival over recent decades in patients with hormone-receptor-positive, HER2-negative breast cancer：a single-center retrospective analysis of prognostic factors. *Jpn J Clin Oncol*. 2018；**48**：248-54.

22）Yamashita H, Ogiya A, Shien T, et al. Clinicopathological factors predicting early and late distant recurrence in estrogen receptor-positive, HER2-negative breast cancer. *Breast Cancer*. 2016；**23**：830-43.

23）Fukuda T, Imai S, Simoda S, et al. Comparison of peripheral nerve block with local infiltration analgesia regarding walking ability after total knee replacement：A retrospective, propensity-score matched-pair cohort study. *J Orthop Surg*（*Hong Kong*）. 2020；**28**：2309499020931656.

24）Iketani R, Imai S, Horiguchi H, et al. Risk stratification for physical morbidity using factors associated with atypical antipsychotic treatment in Parkinson's disease：A retrospective observational study using administrative claims data. *J Clin Neurosci*. 2020；**76**：189-94.

25）Hayashi T, Matsushima M, Bito S, et al. Predictors Associated with Survival Among Elderly In-Patients Who Receive Cardiopulmonary Resuscitation in Japan：An Observational Cohort Study. *J Gen Intern Med*. 2019；**34**：206-10.

原著

日本国内の実臨床における医療機関レセプト・電子カルテデータベースを用いたリアルワールドでの高リスクホルモン受容体陽性/HER2陰性早期乳癌の臨床転帰

――医療機関レセプト・電子カルテデータベースを用いた研究――

井上紀彦[1]，谷　拓朗[1]，金沢奈津子[1]，徳永えり子[2]，青儀健二郎[3]，蔡　志紅[4]，
大佐賀　智[4]，川口　耕[4]，谷澤欣則[4]

1：国立病院機構 本部 総合研究センター 診療情報分析部
2：国立病院機構 九州がんセンター
3：国立病院機構 四国がんセンター
4：日本イーライリリー株式会社 研究開発・メディカルアフェアーズ統括本部

責任著者連絡先：日本イーライリリー株式会社 研究開発・メディカルアフェアーズ統括本部　谷澤欣則
〒651-0086　神戸市中央区磯上通5-1-28
Tel：078-242-9624　E-mail：tanizawa_yoshinori@lilly.com

要　旨

背景：高リスクホルモン受容体陽性（HR＋）/ヒト上皮成長因子受容体2陰性（HER2－）早期乳癌（EBC）患者の臨床転帰および治療パターンに関する本邦のリアルワールドエビデンスは少ない。国立病院機構が構築したデータベースを用いた高リスク集団の最近のエビデンスを提供する。

方法：Ⅰ～Ⅲ期乳癌と診断され2015年4月～2020年3月に乳房切除術を受けた成人女

性を対象とした。手術時DPCデータが「≧N2」または「N1かつ≧T3」の患者を高リスクとした。主要目的は無病生存期間（DFS）とDFSに影響を及ぼす因子の評価，副次目的は患者特性および治療パターンの記述とした。

結果：対象患者8579名中366名が高リスクだった。病期以外の特性は全て高リスク集団と全体集団で同様だった。術前薬物療法（45.4％対7.9％）と術後化学療法（36.6％対19.4％）の使用頻度は高リスク集団の方が全体集団より高かった。最もよく使用された化学療法はシクロホスファミドを含むレジメンだった。5年DFS率は高リスク集団の方が全体集団より低かった（80.4％対95.5％）。高リスクは全体集団で癌の再発/死亡に影響を及ぼす最大の因子だった（ハザード比5.0，95％CI 3.7～7.0，p＜0.001）。

結論：HR＋/HER2－EBC患者のうち，高リスク患者の方が予後不良となる確率が有意に高いことが示された。高リスク集団の治療転帰のさらなる改善が望まれる（**Fig.**：本調査結果の概要）。

Table S1 Definition of variables

Definition of treatment lines with endocrine and chemotherapeutic drugs	
Variables	Definition
Neoadjuvant chemotherapy	
1st neoadjuvant chemotherapy and its start date	The first prescription of any endocrine and chemotherapeutic drug in the 180 days time period before the index date was defined as the start of the 1st neoadjuvant therapy. The combination of all endocrine and chemotherapeutic drugs prescribed within 21 days from the start of the therapy comprises the regimen.
1st neoadjuvant chemotherapy's end date	The date on which all the constituent drugs of 1st neoadjuvant therapy regimen are terminated (i.e. the date of last estimated dose), or one day before any other endocrine and chemotherapeutic drug that was not a part of the regimen was started, or one day before the index date, whichever occurred earliest.
2nd neoadjuvant chemotherapy and its start date	The first prescription of any endocrine and chemotherapeutic drug after the end date of the 1st neoadjuvant therapy and also within 180 days time period before the index date is defined as the start of the 2nd neoadjuvant therapy. The combination of all endocrine and chemotherapeutic drugs prescribed within 21 days from the start of the 2nd neoadjuvant therapy comprises the regimen.
2nd neoadjuvant chemotherapy's end date	The date on which all the constituent drugs of the 2nd neoadjuvant therapy regimen are terminated (i.e. the date of last estimated dose), or one day before any other endocrine and chemotherapeutic drug that was not a part of the regimen was started, or one day before the index date, whichever occurred earliest.
Adjuvant chemotherapy	
1st adjuvant chemotherapy and its start date	Defined by the prescriptions of any chemotherapeutics within 60 days after the index date. The combination of all endocrine and chemotherapeutic drugs prescribed within 21 days from the first prescription (i.e. start date of the therapy) comprises the regimen.
1st adjuvant chemotherapy's end date	The date on which all the constituent drugs of 1st adjuvant chemotherapy regimen are terminated, or one day before any other endocrine and chemotherapeutic drug that was not a part of the regimen was started.
Duration of 1st adjuvant chemotherapy	Interval between the start and end dates of the therapy
2nd adjuvant chemotherapy and its start date	Defined by the prescriptions of any chemotherapeutics within 60 days after the end of the 1st adjuvant chemotherapy. The combination of all endocrine and chemotherapeutic drugs prescribed within 21 days from the first prescription (i.e. start date of the therapy) comprises the regimen.
2nd adjuvant chemotherapy's end date	The date on which all the constituent drugs of the 2nd adjuvant chemotherapy regimen are terminated, or one day before any other endocrine and chemotherapeutic drug that was not a part of the regimen was started.
Duration of 2nd adjuvant chemotherapy	Interval between the start and end dates of the therapy

(continued)

（Table S1　continued）

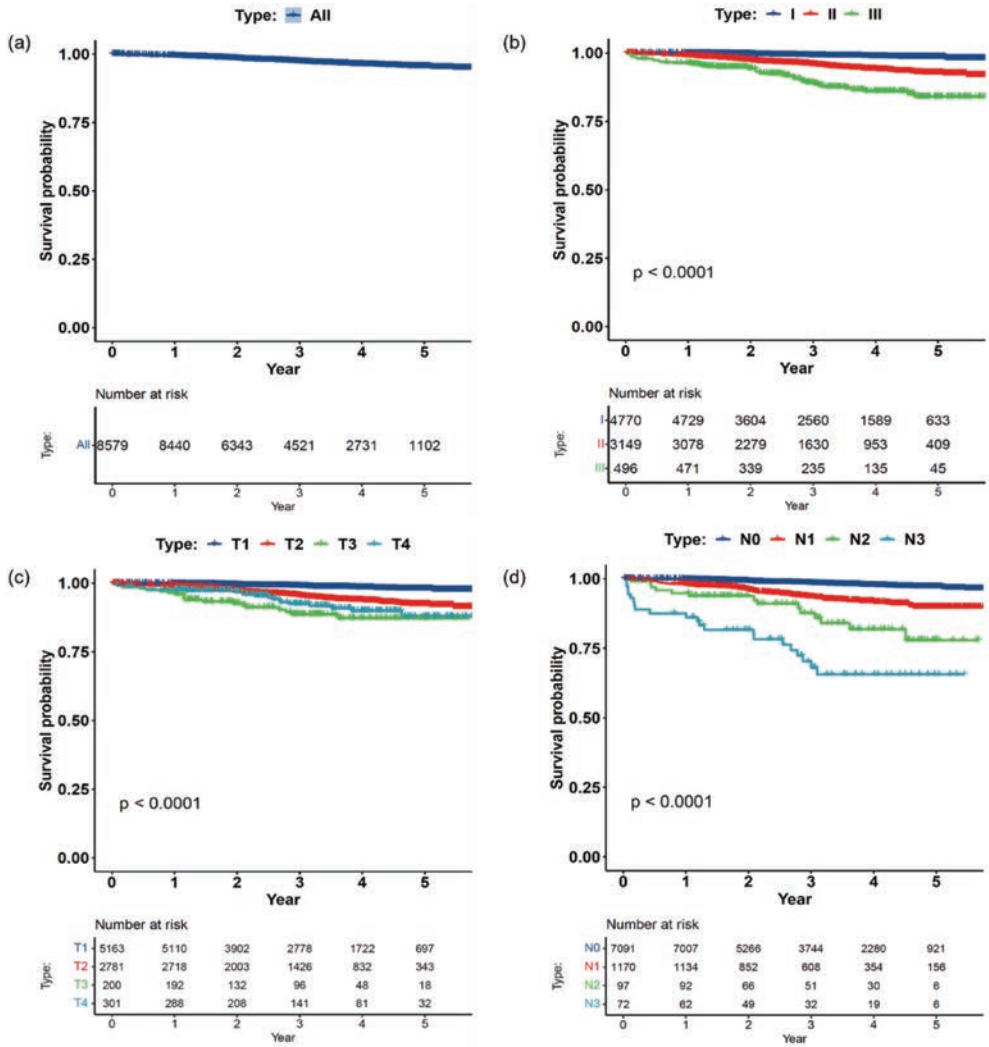
Variables	Definition
Adjuvant ET	
1st adjuvant ET and its start date	Defined by the prescription of any endocrine drug within 60 days after the index date（for patients without adjuvant chemotherapy）or within 60 days after the end date of the adjuvant chemotherapy（for patients with adjuvant chemotherapy）. The combination of all endocrine and chemotherapeutic drugs that were prescribed within 21 days from the first prescription（i.e. start date of the first ET）comprises the regimen.
1st adjuvant ET's end date	The date on which all the constituent drugs of 1st adjuvant ET regimen are terminated, or one day before any other endocrine and chemotherapeutic drug that was not a part of the regimen was started.
Duration of 1st adjuvant ET	Interval between the start and end dates of the therapy
2nd adjuvant ET and its start date	Defined by the prescription of any endocrine drug within 30 days after the end of the 1st adjuvant ET. The combination of all endocrine and chemotherapeutic drugs that were prescribed within 21 days from the first prescription（i.e. start date of the 2nd ET）comprises the regimen.
2nd adjuvant ET's end date	The date on which all the constituent drugs of the 2nd adjuvant ET regimen are terminated, or one day before any other endocrine and chemotherapeutic drug that was not a part of the regimen was started.
Duration of 2nd adjuvant ET	Interval between the start and end dates of the therapy
Variables and codes used for lasso-based and rule-based algorithms for metastasis/recurrence and death events	
Variables	Code
Presence of ICD-10 code C77[a]	C77：Secondary and unspecified malignant neoplasm of lymph nodes
Presence of ICD-10 code C78	C78：Secondary malignant neoplasm of respiratory and digestive organs
Presence of ICD-10 code C79[b]	C79：Secondary malignant neoplasm of other and unspecified sites
Presence of disease codes for recurrent breast cancer	Japanese claims disease code 1749009 "breast cancer recurrence" Japanese claims disease code 8849816 "breast cancer post-operative recurrence on chest wall" Japanese claims disease code 8849815 "breast cancer local recurrence"
Prescription of MBC drugs[c]	Fulvestrant, Denosumab, Bevacizumab, Everolimus, Palbociclib, Abemaciclib, Capecitabine, Gemcitabine, S-1, Nab-Paclitaxel, Irinotecan, Eribulin, Vinorelbine, Olaparib
Death event	Presence of either death record in the discharge summary, death outcome of diagnosis, claims procedures implying death, or claims comment codes implying death

a：The Japanese claims disease code 8842679 for "axillary lymph node metastases" that is mapped to ICD-10 C77 was excluded from this definition.

b：The Japanese claims disease code 8848981 for "metastatic breast cancer" and some other disease codes explicitly indicating metastasis of non-breast primary cancer that are mapped to ICD-10 C79 were excluded from this definition.

c：This variable was defined as true when any of those drugs were prescribed 6 months after the index date or later. This variable was only used for the lasso-based algorithm.

ET：endocrine therapy, ICD：international classification of diseases 10th edition, MBC：metastatic breast cancer

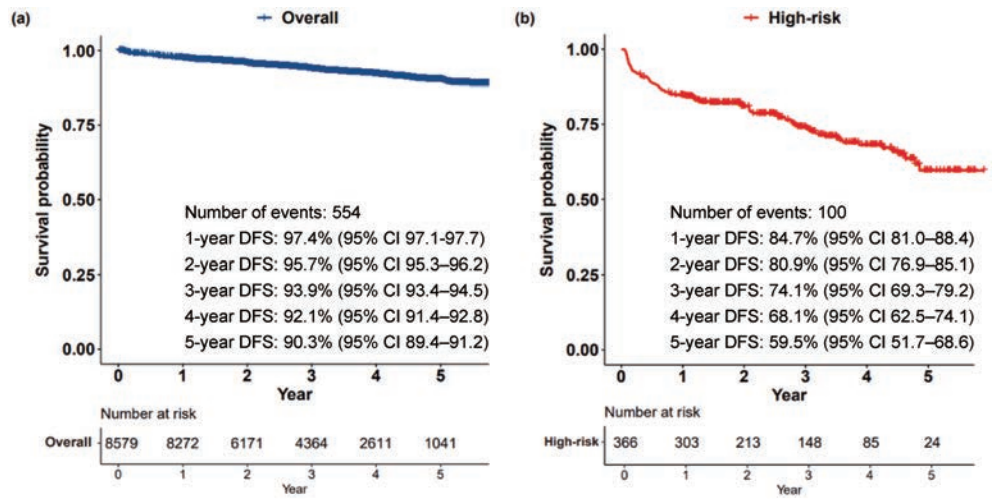


Disease-free survival was analyzed using the Kaplan-Meier method with Log-rank test.

Figs. a-d represent lasso-based disease-free survival of the overall study population, population stratified by stage Ⅰ-Ⅲ, population stratified by T-stage, and population stratified by N-stage, respectively.

EBC：early breast cancer, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive

Fig. S1 Disease-free survival of overall patients with HR＋, HER2－EBC（lasso-based）



Disease-free survival was analyzed using the Kaplan-Meier method with Log-rank test.
Online resources 6a and 6b represent rule-based disease-free survival of the overall and high-risk patients, respectively.

CI：confidence interval, DFS：disease-free survival, EBC：early breast cancer, HER2－：human epidermal growth factor receptor 2 negative, HR＋：hormone receptor-positive

Fig. S2 Disease-free survival of overall patients and high-risk patients with HR＋, HER2－EBC（rule-based）

（Received for publication September 12, 2023）