Biomolecules and Biomedicine
ISSN: 2831-0896 (Print) | ISSN: 2831-090X (Online)
Journal Impact Factor® (2025): 2.8
CiteScore® (2025): 4.7
www.biomolbiomed.com | blog.bjbms.org

SUPPLEMENTAL DATA

Circulatory and other-cause mortality after neoadjuvant versus adjuvant chemotherapy in women with stage II–III breast cancer

Burcu Zihni[1], İsmail Zihni[2*], Mehmet Zafer Sabuncuoğlu[2]

[1]Department of Cardiology, Antalya City Hospital, Antalya, Türkiye;

[2]Department of General Surgery, Akdeniz University Medical Faculty, Antalya, Türkiye.

*Correspondence to **İsmail Zihni**: ismailzihni@gmail.com

Full article is available at the following link: Circulatory and other-cause mortality after neoadjuvant versus adjuvant chemotherapy in women with stage II–III breast cancer

Supplemental file 1. Data provenance

All results were derived from a single case-listing export of 509,208 records obtained from SEER Research Data (17 registries, November 2025 submission), encompassing diagnoses from 2000 to 2023 with follow-up through December 2023. The follow-up duration is verifiable within the export, with a maximum observed follow-up of 13.9 years for women diagnosed in 2010 and 8.9 years for those diagnosed in 2015, which is consistent with a December 2023 cut-off date. All numerical data presented in Figure 1, Tables 1–4, and Tables S1–S12 were regenerated directly from this singular export.

Suplemental file 2. SEER variables and operational definitions

Supplemental file 3. Rationale for excluding diagnosis years 2016–2020. Anatomic AJCC 7th-edition staging data are exclusively available for the period of 2010–2015. Commencing in 2018, SEER provides the Derived EOD-2018 stage group, which deviates from anatomic staging and frequently reclassifies tumors into lower stage categories. For the years 2016–2017, neither of these staging variables is accessible. Empirically, the proportion of women diagnosed with Stage II–III disease consistently ranged from 46.5–46.8% annually between 2010 and 2015. However, this proportion significantly decreased to 20.1–21.7% during 2018–2020. The Combined Summary Stage cannot bridge this data gap, as its localized and regional classifications do not directly correspond to AJCC stage groupings. Consequently, the cohort was restricted to a singular staging era to maintain data consistency and comparability.

Supplemental file 4. Propensity model and weight diagnostics

The propensity model employed a logistic regression to assess the impact of neoadjuvant treatment, considering various covariates: age band, race and origin, marital status, calendar year, clinical stage, molecular subtype, tumor grade, and surgical type. To ensure comprehensive analysis, unknown values, including unspecified molecular subtypes, were maintained as distinct categories, thereby preventing the exclusion of any participant. Stabilized weights were subsequently trimmed at the 1st and 99th percentiles.

Supplemental file 5. Conditional estimand and bootstrap procedure

The conditional analysis is an unweighted, descriptive assessment of natural history within an age stratum, not a treatment contrast. A binary crossover rule is not employed due to its reliance on an arbitrary threshold and the inherent instability of the tail of follow-up data. The estimand is defined as the difference in cumulative incidence between circulatory and breast cancer among women alive at the landmark, evaluated at 0.25-year intervals. Seven years post-landmark was pre-specified as the primary time point because fewer than 10% of the stratum remained at risk thereafter; the 8-year value serves as an unstable sensitivity analysis. Confidence intervals are percentile intervals derived from 1,000 individual-level bootstrap resamples.

Supplemental file 6. Sensitivity analyses

Supplemental file 7. Reproducibility

All analyses were performed in Python and independently reproduced in R. Agreement was required within 1% on the log-hazard scale for all hazard ratios and within one percentage point for all cumulative-incidence estimates. No models incorporated penalization.

Supplemental file 8. Sensitivity of the primary analysis to the exclusion of surgical procedure type from the propensity model

The surgical procedure type, as recorded in SEER, represents the definitive procedure actually performed. In cases where patients received preoperative treatment, this procedure is selected after treatment administration and may be influenced by treatment response. Consequently, it can be considered a post-treatment variable rather than a pretreatment confounder. Conditioning on such a variable could potentially lead to overadjustment or collider bias. Nevertheless, surgical procedure type was incorporated into the primary propensity model due to its significant influence on both the decision for preoperative treatment and subsequent radiotherapy. Furthermore, its exclusion resulted in a notable pretreatment imbalance (standardized difference of 0.223 for breast conservation versus mastectomy). To confirm that the primary conclusion remains robust irrespective of this modeling choice, the entire analysis was re-estimated with surgical procedure type excluded from the propensity model, while all other specifications remained constant.

Table S1. SEER*Stat variables utilized

Variable	Years	Role
Site recode ICD-O-3/WHO 2008	all	Restriction to breast
Sex	all	Restriction to women
Diagnostic confirmation	all	Microscopically confirmed disease
Sequence number	all	First malignant primary only
Combined summary stage with expanded regional codes	2004+	Exclusion of distant disease
Derived AJCC stage group, 7th ed	2010–2015	Anatomic clinical stage II–III (cohort definition)
Derived EOD 2018 stage group	2018+	Not used; basis for excluding 2016–

recode		2020
Chemotherapy recode	2004+	Required to be Yes in both groups
RX Summ–Systemic/Sur Seq	2007+	Exposure: recorded systemic-therapy/surgery sequence
RX Summ–Surg Prim Site	1998–2023	Definitive breast surgery (Table S2)
Breast subtype	2010+	Hormone-receptor/HER2 subtype; unknown retained as a category
Grade recode (thru 2017)	–2017	Grade covariate
SEER cause-specific death classification	all	Breast-cancer death
COD to site recode	all	Circulatory and other-cause death
Vital status recode	all	Censoring
Survival months	all	Time scale

Abbreviations: AJCC, American Joint Committee on Cancer; COD, cause of death; EOD, Extent of Disease; HER2, human epidermal growth factor receptor 2; ICD-O-3, International Classification of Diseases for Oncology, Third Edition; SEER, Surveillance, Epidemiology, and End Results; WHO, World Health Organization.

Table S2. Classification of RX Summ–Surg Prim Site codes

Code range	Procedure	Classification
00	No surgery of primary site	Excluded
09–19	Local tumour destruction or excisional biopsy	Excluded (diagnostic)
20–24	Partial mastectomy or lumpectomy	Definitive – breast conservation
30–80	Subcutaneous, total, modified radical and radical mastectomy	Definitive – mastectomy

Code range	Procedure	Classification
90, 99	Surgery not otherwise specified or unknown	Excluded

Diagnostic procedures are excluded as definitive surgical interventions.

Table S3. Proportion of stage II–III disease by diagnosis year and staging system

Diagnosis year	Staging variable	Women screened	Stage II–III	%
2010	AJCC 7th	41,784	19,527	46.7
2011	AJCC 7th	43,596	20,302	46.6
2012	AJCC 7th	44,109	20,513	46.5
2013	AJCC 7th	44,893	20,957	46.7
2014	AJCC 7th	45,719	21,309	46.6
2015	AJCC 7th	47,529	22,244	46.8
2016–2017	none available	95,977	—	—
2018	EOD 2018	49,248	10,376	21.1
2019	EOD 2018	50,844	10,198	20.1
2020	EOD 2018	45,509	9,894	21.7

Abbreviations: AJCC, American Joint Committee on Cancer; EOD, Extent of Disease.

Table S4. Classification of exposure based on RX summary-systemic/sequence

Recorded value	Classification
Systemic therapy before surgery	Neoadjuvant
Systemic therapy both before and after surgery	Neoadjuvant
Systemic therapy after surgery	Adjuvant
Intraoperative systemic therapy	Excluded
Intraoperative systemic therapy with other systemic therapy before or after surgery	Excluded
Surgery both before and after systemic therapy	Excluded
No systemic therapy and/or surgical procedures	Excluded
Sequence unknown or blank	Excluded

Additionally, a positive chemotherapy indicator was required in both groups.

Table S5. Cause-of-death hierarchy

Order	Category	Definition
1	Breast-cancer death	SEER cause-specific death classification: dead, attributable to this cancer
2	Circulatory death	Among remaining deaths: diseases of the heart; hypertension; cerebrovascular disease; atherosclerosis; aortic aneurysm and dissection; other arterial disease (ICD-10 I00–I99)
3	Other-cause death	All remaining deaths
0	Censored	Alive at last contact

Each death is assigned to a single, mutually exclusive category, ordered hierarchically. The restricted cardiac endpoint (sensitivity analysis) encompasses diseases of the heart only (ICD-10 I00–I09, I11, I13, I20–I51). **Abbreviations**: ICD-10, International

Classification of Diseases, Tenth Revision; SEER, Surveillance, Epidemiology, and End Results.

Table S6. Proportion of unknown molecular subtype by age at diagnosis

Age at diagnosis	*n*	Subtype unknown
Under 50 years	26,913	3.5%
50–64 years	32,329	4.1%
65–74 years	11,952	4.4%
75 years or older	3,039	4.7%

Of the 1,996 women aged 75 years or older who were alive at the landmark, 19.1% had received neoadjuvant treatment. Among these women, 47.7% presented with hormone-receptor-positive/HER2-negative disease, 32.6% with HER2-positive disease, and 19.7% with triple-negative disease. It is important to note that this cohort may not be representative of the general population of older women diagnosed with breast cancer. **Abbreviation**: HER2, human epidermal growth factor receptor 2.

Table S7. Impact of excluding surgical modality from the propensity model

Quantity	Primary model (type of surgery included)	Sensitivity model (type of surgery removed)
Maximum standardized difference after weighting	0.021	0.022
Effective sample size (of 74,233)	68,073	68,284
Weighted 10-year cumulative incidence,	2.12% vs 2.14%	2.12% vs 2.14%

neoadjuvant vs adjuvant		
Cause-specific hazard ratio for circulatory death (95% CI)	1.087 (0.963-1.226)	1.095 (0.971-1.235)

Both columns originate from a single, independent refit of the entire pipeline, ensuring their direct comparability. This refit closely, though not identically, reproduces the primary analysis. Consequently, its primary column exhibits marginal differences from Table 3 and Table S8 (cause-specific hazard ratio: 1.087 versus 1.092; effective sample size: 68,073 versus 68,357; 10-year cumulative incidence: 2.12% and 2.14% versus 2.06% and 2.10%). This table primarily establishes a within-run comparison. Removing the "type of surgery" variable maintains balance, leaves the weighted 10-year cumulative incidence and risk difference unchanged, and shifts the cause-specific hazard ratio by 0.008. Therefore, the primary conclusion is not attributable to conditioning on a post-treatment variable. **Abbreviation**: CI, confidence interval.

Table S8. Weight diagnostics (*n* = 74,233)

Diagnostic	Value
Mean weight	0.996
Standard deviation	0.292
Median	0.934
Maximum after trimming	2.231
Effective sample size	68,357 (of 74,233)
Maximum standardized difference before weighting	0.365
Maximum standardized difference after weighting	0.022

Table S9. Sensitivity of the primary association to landmark definition and subgroup restriction

Analysis	*n*	Circulatory deaths	Deaths excluded by the landmark	Estimate (95% CI)
3-month landmark	74,442	1,646	43	1.082 (0.959–1.219)
6-month landmark (primary)	74,233	1,625	252	1.092 (0.968–1.231)
9-month landmark	73,875	1,600	610	1.103 (0.977–1.245)
12-month landmark	73,422	1,581	1,063	1.103 (0.976–1.246)
Hormone-receptor-negative subgroup	18,508	459	—	1.119 (0.912–1.374)
Triple-negative (secondary)	13,121	325	—	1.177 (0.920–1.505)
HER2-positive subgroup	16,919	365	—	0.994 (0.784–1.259)

Point estimates for the primary analyses are presented in Table 3. This table also details the analytic samples and event counts that underpin these estimates, which vary with changes in the landmark. The number of women and circulatory deaths consistently decrease as the landmark lengthens, as later landmarks inherently exclude women who experienced early mortality; the estimate itself remains constant. The row designated "HER2-positive" facilitates a comparison of recorded sequences exclusively within the HER2-positive subtype. It is important to note that HER2-directed therapy receipt is not documented in SEER, thus this comparison does not reflect trastuzumab exposure versus non-exposure. Given that the primary estimate is null, an E-value is not reported, as such metrics offer no informative value when the confidence interval encompasses the null hypothesis. **Abbreviations**: CI, confidence

interval; HER2, human epidermal growth factor receptor 2; SEER, Surveillance, Epidemiology, and End Results.

Table S10. Leading categories of other-cause deaths

Category	Whole cohort (*n* = 4,441)	75 y or older, alive at landmark (*n* = 423)
Second primary cancers (total)	1,123 (25.3%)	73 (17.3%)
Lung and bronchus	274	22
Pancreas	122	9
Colon excluding rectum	82	9
Acute myeloid leukaemia	85	4
Other cancer sites	560	29
Non-cancer causes (total)	3,318 (74.7%)	350 (82.7%)
Other or ill-defined cause	1,293	161
Alzheimer disease or dementia	175	53
Chronic obstructive pulmonary disease	321	33
Death certificate unavailable	401	30
Accidents	218	16
Pneumonia and influenza	128	12
Nephritis or renal disease	109	12
Diabetes mellitus	177	10
Septicaemia	170	9
Other listed non-cancer causes	326	14

Cancer deaths were defined as any malignant site other than breast. In the oldest stratum, the largest single category was not a specific disease: 161 deaths (38% of

other-cause deaths) were coded as "other" or "ill-defined," and an additional 30 (7%) lacked a usable death certificate. Among specifically coded deaths, dementia was the leading cause. This uncertainty in cause assignment for the very old suggests a similar degree of uncertainty in assigning circulatory causes within the same cohort.

Table S11. Risk set composition: number of individuals at risk, interval event counts, and calendar time for women aged 75 or older at diagnosis who were alive at the five-year landmark (*n* = 1,996)

Years after landmark	At risk	Circulatory deaths	Breast-cancer deaths	Diagnosis years contributing
0	1,996	35	49	2010–2015
1	1,831	29	55	2010–2015
2	1,659	36	38	2010–2015
3	1,378	38	36	2010–2015
4	991	36	14	2010–2014
5	654	21	10	2010–2013
6	409	9	4	2010–2012
7 (primary)	208	10	3	2010–2011
8 (unstable)	57	2	0	2010

A woman diagnosed in 2015 can contribute a maximum of 3.4 years of data following the landmark, while a woman diagnosed in 2013 can contribute a maximum of 5.4 years. Consequently, by the primary time point, the risk set is restricted to women diagnosed between 2010 and 2011. Therefore, estimates at and beyond seven years are partially confounded with the calendar period.

Table S12. Difference in cumulative incidence (percentage points) of circulatory versus breast cancer among women aged 75 years or older at diagnosis and alive at the five-year landmark

Analysis	3 y	5 y	7 y (primary)	8 y (unstable)
Broad circulatory (I00–I99), *n* = 1,996	−2.11 (−3.63, −0.57)	−0.25 (−2.21, +1.97)	+1.88 (−0.81, +4.79)	+4.13 (+0.38, +7.72)
Heart disease only, *n* = 1,996	−3.68 (−5.18, −2.17)	−3.26 (−5.23, −1.20)	−2.01 (−4.40, +0.66)	−0.43 (−3.86, +2.75)
Diagnosed 2010–2012, *n* = 987	−2.34 (−4.58, −0.09)	−1.41 (−4.16, +1.38)	+0.80 (−2.38, +3.97)	+3.04 (−0.79, +7.02)

Table S13. Timing of post-landmark deaths in women aged 75 years or older at diagnosis

Cause	Deaths	Median (years)	Interquartile range
Breast cancer	209	2.00	1.00–3.50
Heart disease	156	3.50	1.73–4.69
Non-cardiac circulatory	60	2.88	1.31–4.19
Other causes	423	2.83	1.33–4.29

Heart disease deaths occurred later than non-cardiac circulatory deaths (Mann-Whitney $p = 0.23$), contrary to predictions for a cerebrovascular-driven transition. The disappearance of this difference under the restricted endpoint reflects the loss of 60 of 216 events, rather than a differently timed process. Breast cancer deaths were notably earlier than circulatory deaths ($p < 0.001$), indicating that the convergence of these curves reflects the exhaustion of breast cancer risk, not an acceleration of circulatory risk.

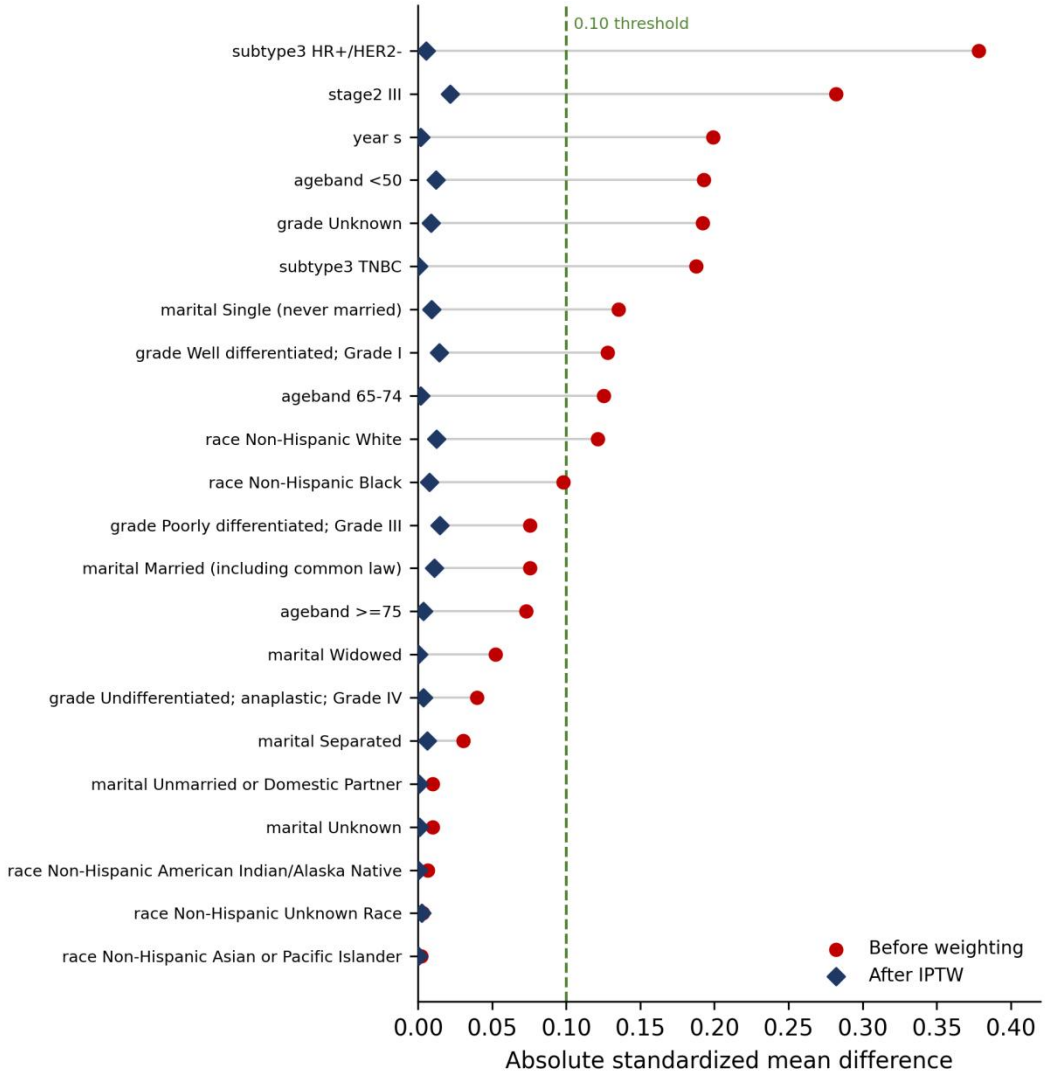


Figure S1. Covariate balance before and after IPTW. Love plot showing the absolute SMDs for covariates included in the propensity model before weighting (red circles) and after IPTW (blue diamonds) between the neoadjuvant and adjuvant sequence groups. Covariates are ordered by decreasing pre-weighting imbalance. The vertical dashed line denotes the reference threshold of 0.10. After IPTW, all absolute SMDs were below 0.10, indicating adequate covariate balance between the groups. **Abbreviations**: HER2, human epidermal growth factor receptor 2; HR, hormone receptor; IPTW, inverse probability of treatment weighting; SMD, standardized mean difference; TNBC, triple-negative breast cancer.

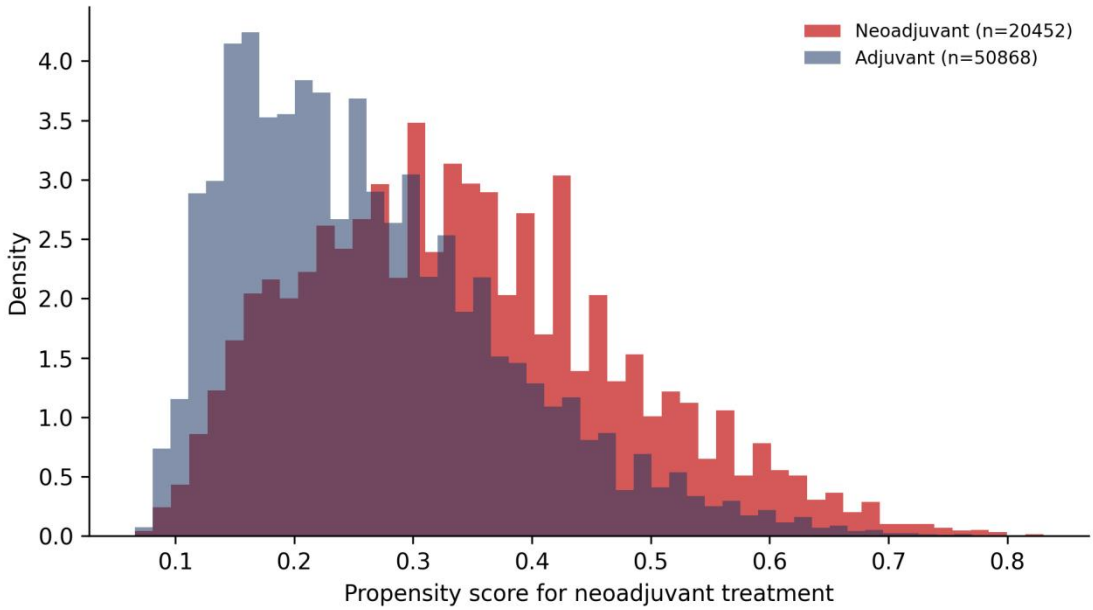


Figure S2. Propensity-score distributions by treatment-sequence group. Overlaid histograms show the distribution of propensity scores for neoadjuvant treatment in the neoadjuvant and adjuvant sequence groups after trimming. The two distributions overlap across the observed score range, indicating adequate common support between the groups. This overlap was also present within the ≥75-year stratum, supporting the practical plausibility of the positivity assumption.

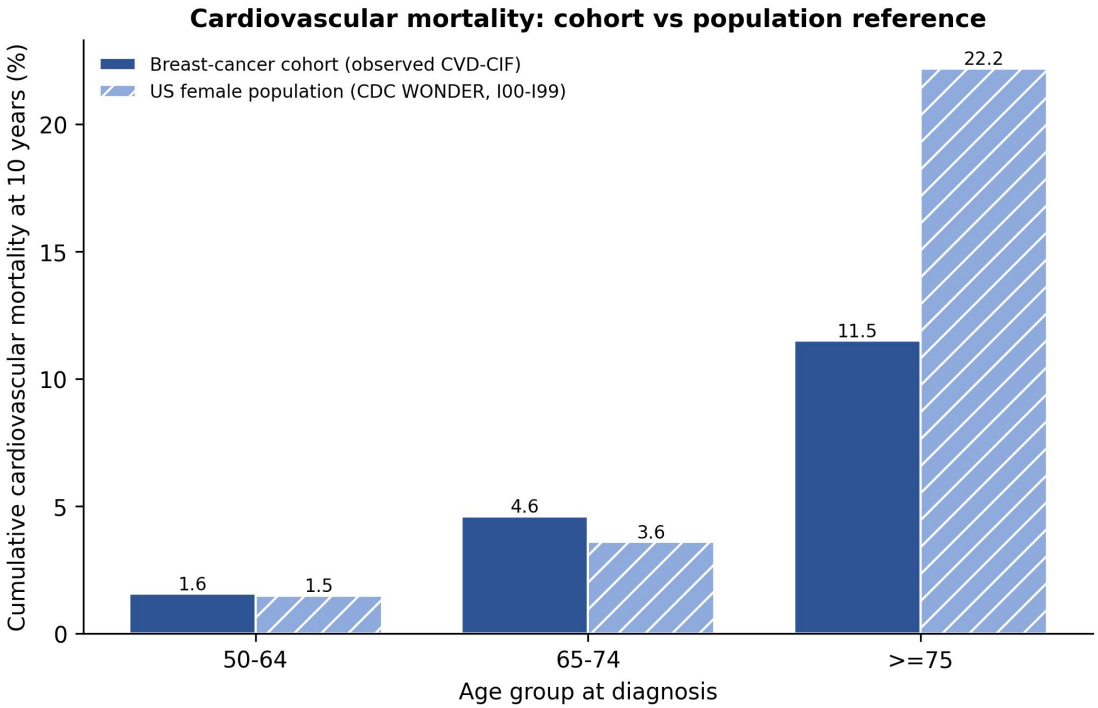


Figure S3. Descriptive comparison of cohort CVD-CIF and US population circulatory mortality by age at diagnosis. Bars compare the observed 10-year CVD-CIF in the breast-cancer cohort with the corresponding female-population circulatory-mortality reference derived from CDC WONDER (ICD-10 codes I00–I99), stratified by age group at diagnosis. This comparison is exploratory and descriptive only and should not be interpreted as a standardized mortality ratio. The cohort estimate is a competing-risk CIF and is therefore depleted by prior BC deaths, particularly in the ≥75-year group, whereas the population reference approximates cumulative circulatory risk from annual female mortality rates; in the ≥75-year category, the population comparator also includes women aged 85 years or older. Because age is grouped categorically rather than by attained age, formal attained-age standardization is not possible. **Abbreviations**: BC, breast cancer; CDC, Centers for Disease Control and Prevention; CIF, cumulative incidence function; CVD, cardiovascular disease; ICD-10, International Classification of Diseases, Tenth Revision.