SUPPLEMENTARY DATA

SUPPLEMENTARY METHODS

CA125 measurements

Serum concentrations of CA125 were determined using a commercially available chemiluminescent microparticle immunoassay (Alinity i CA125 Reagent, Abbott Diagnostics). The intraassay coefficients of variation were 4.7% for concentrations < 40 U/mL and 3.6% for concentrations > 40 U/mL. The validated analytical range spanned from 1.1 to 1000 U/mL.

Outcome adjudication

All outcomes were ascertained and adjudicated retrospectively. Site investigators systematically reviewed electronic medical records from the SIA-GAIA and Orion Clinic databases, which comprehensively document all health care encounters within the Valencian Community public health care system, including emergency department visits, hospitalizations, and outpatient care. CA125 values were masked in the electronic case report forms during the adjudication process.

Statistical analysis

Descriptive statistics

Continuous variables are summarized as mean ± standard deviation (SD) or median (interquartile range [IQR]) according to their distributional properties; categorical variables are reported as counts and percentages. Baseline characteristics were compared across CA125 categories using a threshold of 35 U/mL, consistent with previously reported prognostic thresholds in heart failure (HF) populations, using the *t* test or Mann–Whitney U test for continuous variables and the chi-square or Fisher exact test for categorical variables, as appropriate.

CA125 predicted trajectories

The predicted CA125 trajectories over time were estimated by joint modeling of longitudinal and survival data (JM), which is a multicomponent approach with 2 submodels (longitudinal and survival). The mortality submodel—with baseline hazard modeled with Weibull parametric distribution—was linked to the longitudinal CA125 submodel via the "current value association" framework. In this approach, the predicted CA125 trajectory is simultaneously adjusted by informative dropout due to patient death. Baseline covariates included age, sex, history of previous HF admissions, ischemic etiology, heart rate (beats/min), systolic blood pressure (mmHg), left ventricular ejection fraction (LVEF, %), hemoglobin (g/dL), estimated glomerular filtration rate (eGFR), NT-proBNP (pg/mL), and use of beta-blockers, angiotensin-converting enzyme inhibitors (ACEIs), angiotensin receptor blockers (ARBs), and sodium-glucose cotransporter 2 (SGLT2) inhibitors. Results are illustrated by plotting the predicted CA125 trajectories (with 95% confidence intervals [CIs]) against time.

CA125 prognostic profile

To visualize the accumulation of risk for HF readmission over time, the Nelson–Aalen method was used to estimate the cumulative hazard function. Separate Nelson–Aalen curves were generated for subgroups defined by baseline CA125 cutpoints or quartiles to compare risk trajectories. To assess the impact of longitudinal CA125 on rehospitalizations, a joint frailty model (JFM) was implemented.[1] In this framework, any HF hospitalization occurring after the index admission was treated as a recurrent event, with all-cause mortality modeled as a terminal event. The recurrent event submodel was fitted using a Royston–Parmar parametric approach, whereas the mortality submodel employed a Weibull baseline hazard; both submodels were jointly modeled through a shared random effect.[1] Results are reported as hazard ratios (HRs) with 95%CIs. The same baseline covariates were incorporated into both submodels of the JFM. In this counting process data structure, each patient's follow-up was partitioned into intervals delimited by measurement times, with CA125 taken as the value recorded at the start of each interval. This ensured that the biomarker level used to estimate

the hazard at any given time was always measured before any event occurring within that interval, thereby avoiding temporal circularity between exposure and outcome.[1]

Stepwise analytical approach

CA125 was log-transformed and first introduced as a main effect. Then, to identify possible time-dependent relationships, we included the interaction terms between log(CA125) and time. Time in both submodels was initially modeled with restricted cubic splines with 3 knots to accommodate nonlinear patterns during the follow-up period. If the omnibus *P* value for interaction was less than .05, the interaction term was dropped from the model and reported instead of the effect of CA125 on that endpoint as main term. For log(CA125), we then reported the HR, 95%CI, and *P* value as a constant effect during the follow-up; we also reported the association between CA125 (back-transformed to its original values) with clinical outcomes using a figure plotting the HR trajectory against time.

Sensitivity analysis

Four additional sensitivity analyses were performed to prove the robustness of the primary JFM. First, to rule out the possibility that CA125 values obtained during rehospitalization episodes inflated the observed associations, all inhospital CA125 measurements were replaced with the last available outpatient or baseline value. Second, to evaluate whether the CA125–outcome association differed by measurement availability, we stratified patients into those with a single inhospital measurement (Group A) and those with at least one postdischarge measurement (Group B), replicating the JFM in each subgroup. Third, the analysis was restricted to patients with de novo acute HF (AHF) (ie, excluding those with a history of prior HF admissions) to verify that the CA125–outcome association was not confounded by recurrent admission phenotypes. Fourth, the cohort was stratified by enrollment period (2008-2015 vs 2016-2023), and the JFM was estimated within each stratum to evaluate the temporal stability of the prognostic signal across eras of evolving background therapy.

As a complementary approach, we used linear mixed-effects models with random patient-level intercepts and restricted maximum likelihood estimation to assess the longitudinal trajectories of log(CA125) over time before the occurrence of the first HF rehospitalization. Log(CA125) was plotted against time, defined as the number of days before the outcome of interest or the end of follow-up. Patients who experienced an event were compared with a control population who remained alive and free of hospitalization during follow-up. For each comparison, 2 *P* values are reported: 1 testing whether the trajectories over time are identical (superimposable), and 1 testing whether they differ in shape over time (nonparallel trajectories, ie, group-by-time interaction).

Finally, we modeled all-cause mortality using flexible parametric survival models (stpm3) with CA125 summarized as the mean during follow-up and categorized into quartiles. Covariate-adjusted standardized survival and restricted mean survival time (RMST) up to 180 days were then estimated using "standsurv", providing contrasts among CA125 quartiles.[2] This analysis was performed as a sensitivity check to complement the joint modeling results, allowing comparison of findings using a simpler survival modeling framework that does not explicitly account for informative dropout.

Ethical considerations

The study was conducted in accordance with the principles outlined in the Declaration of Helsinki and was approved by the local institutional review board. Given its retrospective, noninterventional design based on routinely collected clinical data, the ethics committee granted a waiver of individual informed consent, in accordance with applicable regulations.

SUPPLEMENTARY RESULTS

CA125 trajectory after discharge

When CA125 trajectories were stratified according to baseline CA125 status, 2 distinct patterns emerged. Patients with elevated baseline CA125 (>35 U/mL) mirrored the overall cohort trajectory, although they started at higher levels (approximately 85 U/mL). After an early subtle peak, these

patients experienced a similar steep initial decline followed by a gradual decrease thereafter (figure 2a of the supplementary data). In contrast, patients with normal baseline CA125 (≤35 U/mL) displayed a subtle early peak but subsequently maintained consistently low and stable levels (approximately 15-20 U/mL) throughout the entire follow-up period (figure 2a of the supplementary data). The joint model-predicted trajectory stratified by baseline CA125 was also consistent with the observed patterns (figure 2b of the supplementary data). These stratified trajectories show that elevated CA125 levels progressively normalized in patients initially presenting high values, with most of this decline occurring within the first 30 days postdischarge.

Sensitivity analysis

Robustness to inhospital CA125 measurements and measurement availability

Replacing all inhospital CA125 measurements with the last available outpatient value yielded HRs nearly identical to the primary analysis (HHF: HR, 1.249; 95%CI, 1.154-1.351; mortality: HR, 1.678; 95%CI, 1.443-1.952; both $P < .001$). When stratified by CA125 measurement availability, Group A (single measurement; n = 2449) had only 10 HHF events (event per variable = 0.67), precluding recurrent-event modeling; a mortality-only Royston–Parmar model yielded a HR of 1.375 (95%CI, 1.246-1.517; $P < .001$). In Group B (≥2 measurements; n = 1531; 4526 observations), the full JFM confirmed associations for both HHF (HR, 1.273; 95%CI, 1.180-1.374; $P < .001$) and mortality (HR, 2.280; 95%CI, 1.726-3.011; $P < .001$). The HHF point estimate in Group B was consistent with the primary analysis, indicating that differential measurement frequency did not inflate the observed associations (table 1 of the supplementary data).

Restriction to de novo AHF and period-stratified analysis

Restricting the analysis to de novo AHF patients (n = 2991) obtained directionally consistent results (HHF: HR, 1.260; 95%CI, 1.148-1.383; mortality: HR, 1.617; 95%CI, 1.331-1.965; both $P < .001$). Similarly, period-stratified analyses (2008-2015 vs 2016-2023) demonstrated that the prognostic

association of longitudinal CA125 was significant in both the 2008 to 2015 stratum (HHF: HR, 1.182; 95%CI, 1.078-1.296; mortality: HR, 1.568; 95%CI, 1.311-1.875) and the 2016 to 2023 stratum (HHF: HR, 1.364; 95%CI, 1.183-1.571; mortality: HR, 1.937; 95%CI, 1.441-2.604) (all $P < .001$).

All-cause mortality modeled using flexible parametric survival models

With CA125 summarized as the mean during follow-up and categorized into quartiles, contrasts of the RMST over 180 days using the lowest CA125 quartile (Q1) as the reference showed a graded loss of event-free days with increasing CA125 levels (figure 6 of the supplementary data). Compared with Q1, patients in Q2 experienced, on average, approximately 5 fewer survival days, while those in Q3 experienced around 8 to 10 fewer days and those in Q4 experienced about 20 fewer days by the end of follow-up (figure 6 of the supplementary data). Similarly, contrasts of standardized survival curves using the lowest CA125 quartile (Q1) as reference showed a clear, graded excess risk associated with higher quartiles (figure 7 of the supplementary data). Compared with Q1, patients in Q2 had modest but progressively widening reductions in survival probability over 180 days. The gap was larger for Q3 and most pronounced for Q4, where survival probability was consistently lower than in Q1, with differences approaching 10% to 15% by the end of follow-up (figure 7 of the supplementary data). These results emphasize the clinically interpretable impact of higher CA125: patients with elevated levels spent substantially fewer days alive and event-free within 6 months.

Table 1 of the supplementary data. Summary of sensitivity analyses

Analysis	**n**	**HHF HR (95%CI)**	**Mortality HR (95%CI)**
Primary analysis	3980	1.239 (1.147-1.339)	1.668 (1.434-1.940)
SA1: replacing inhospital (HHF event) CA125 values with last outpatient measurements	3980	1.249 (1.154-1.351)	1.678 (1.443-1.952)
SA2: De novo AHF only	2991	1.260 (1.148-1.383)	1.617 (1.331-1.965)
SA3a: 2008-2015 period	2529	1.182 (1.078-1.296)	1.568 (1.311-1.875)
SA3b: 2016-2023 period	1451	1.364 (1.183-1.571)	1.937 (1.441-2.604)
SA4a: Group A (single measurement)†	2449	—	1.375 (1.246-1.517)
SA4b: Group B (multiple measurements)	1531	1.273 (1.180-1.374)	2.280 (1.726-3.011)

All *P* values are < .001. HRs represent the effect of a 1-log-unit increase in CA125 on each endpoint.

†Group A: only 10 HHF events (events per variable = 0.67 for 15 covariates), which precluded recurrent-event modeling; a mortality-only model is therefore reported (Royston–Parmar, df = 3).

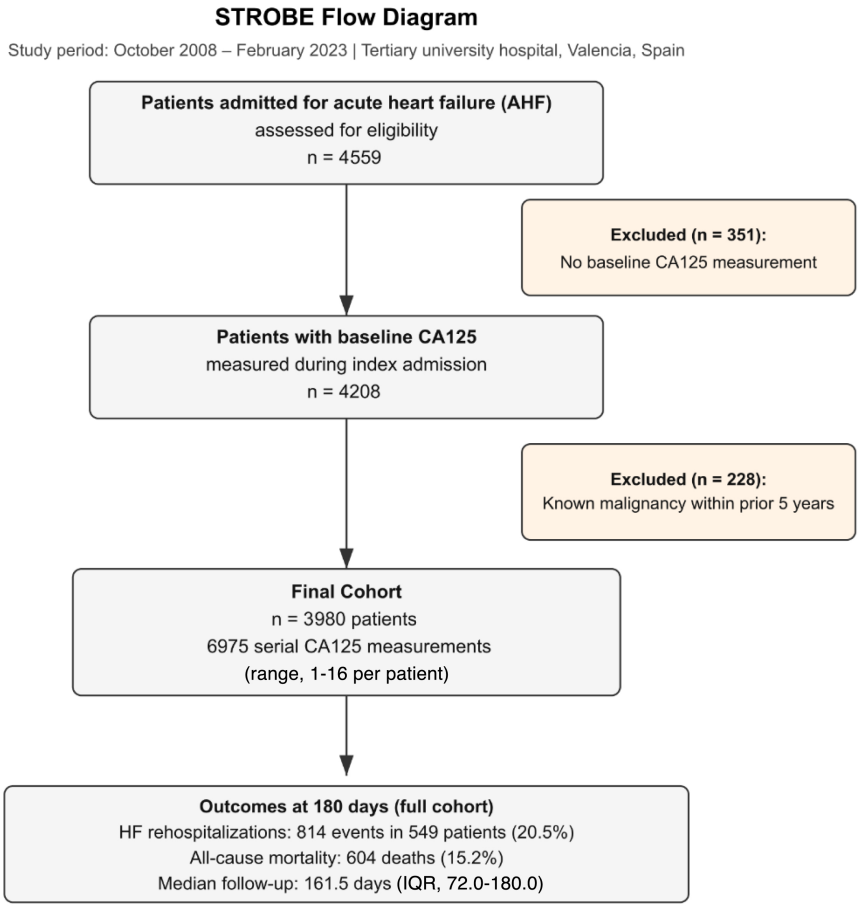
AHF, acute heart failure; CI, confidence interval; HHF, heart failure rehospitalization; HR, hazard ratio; LOCF, last observation carried forward; SA, sensitivity analysis.

Table 2 of the supplementary data. Trajectory of log(CA125) levels before a first HF rehospitalization

Time period	Predicted mean log(CA125) (95%CI)	*P* value
Remained alive and free of HF rehospitalization after the index event		
−180 to −91 d	4.22 (4.14-4.29)	< .001
−90 to −61 d	3.94 (3.89-4.03)	< .001
−60 to −31 d	3.83 (3.74-3.92)	< .001
−30 to 0.1 d	3.50 (3.41-3.60)	< .001
0 d	3.19 (3.12-3.25)	< .001
Experienced a HF rehospitalization after the index event		
−180 to −91 d	4.21 (4.11-4.30)	< .001
−90 to −61 d	4.14 (4.03-4.26)	< .001
−60 to −31 d	4.14 (4.04-4.23)	< .001
−30 to 0.1 d	4.15 (4.05-4.24)	< .001
0 d	4.25 (4.17-4.32)	< .001

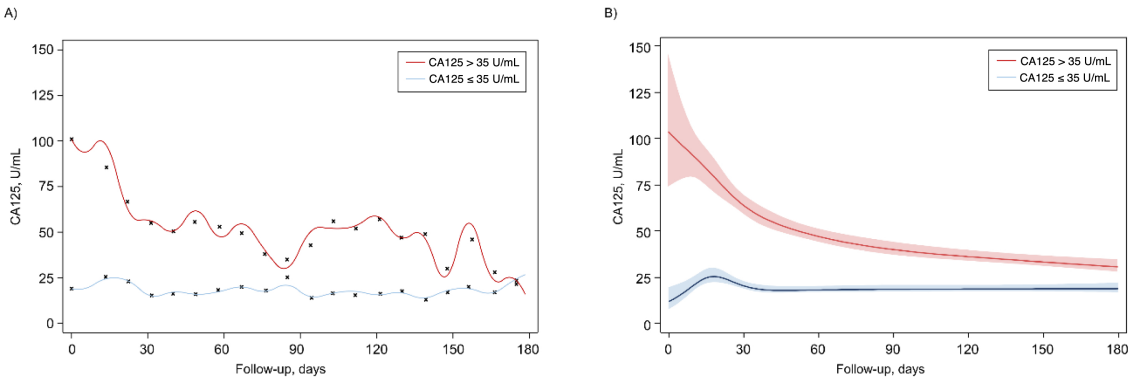
CA125, serum antigen carbohydrate 125; CI, confidence interval; HF, heart failure; Log, logarithmic transformation.

Figure 1 of the supplementary data. Flow chart.



AHF, acute heart failure; CA125, carbohydrate antigen 125; HF, heart failure; IQR, interquartile range

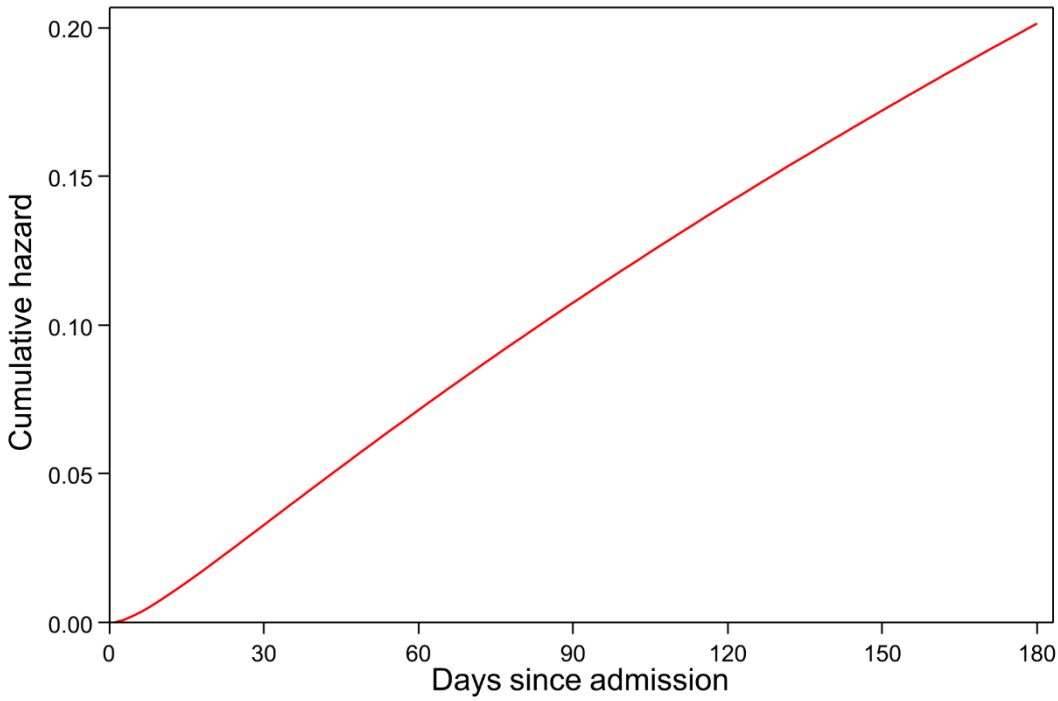
Figure 2 of the supplementary data. CA125 trajectory after discharge stratified according to the baseline CA125 value (> 35 vs ≤ 35 U/mL).



(A) Unadjusted median CA125 levels during 180 days of follow-up, shown separately for patients with baseline CA125 ≤ 35 U/mL (blue) and > 35 U/mL (red). Dots represent observed medians at different time points, and lines depict smoothed trajectories. (B) Adjusted CA125 trajectories estimated from joint longitudinal–survival models accounting for death as a competing risk. The solid lines represent the average predicted course of CA125 within each subgroup, with shaded areas denoting 95% confidence intervals. In both strata, the spline-defined time effect was highly significant (global Wald test $P < .001$), indicating a nonlinear trajectory of CA125 levels during follow-up.

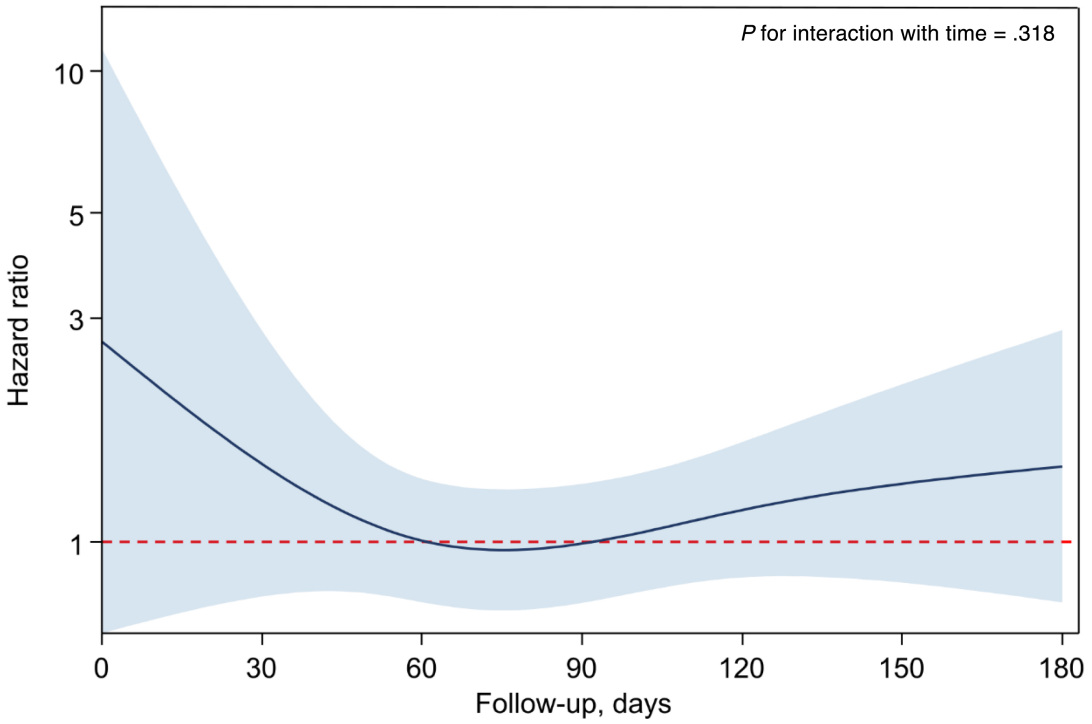
CA125, carbohydrate antigen 125.

Figure 3 of the supplementary data. Nelson–Aalen cumulative hazard function estimated from a multilevel lognormal survival model for recurrent HF hospitalizations.



The Y-axis values represent the integral (accumulation) of the instantaneous hazard rate during the follow-up period and are interpreted as the expected cumulative number of events per patient, assuming continuous accumulation of risk. For example, a cumulative hazard of 0.20 at 180 days implies roughly 0.2 events per patient during the 180-day follow-up, assuming a constant hazard throughout the interval (interpreted as approximately 20 events per 100 patients). The cumulative hazard increased steadily during the 180-day follow-up, indicating a continuous accumulation of rehospitalization risk in the study cohort.

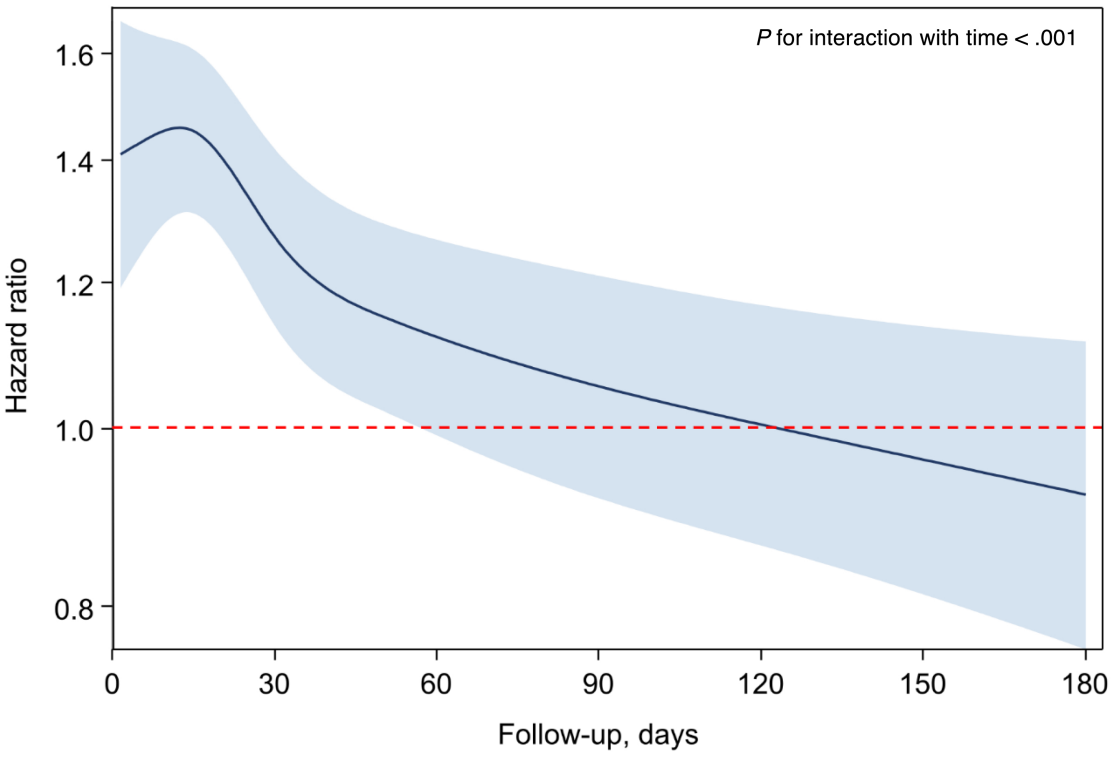
Figure 4 of the supplementary data. Time-varying hazard ratio for recurrent HF hospitalizations.



The solid line represents the estimated hazard ratio for recurrent HF hospitalization associated with a 10 U/mL increase in time-updated CA125, and the shaded band denotes the corresponding 95% confidence interval. Hazard ratios are displayed on a logarithmic scale and referenced to HR = 1. The association remained broadly stable over time, and there was no statistical evidence of a time-dependent effect (*P* for interaction with time = .318).

CA125, carbohydrate antigen 125.

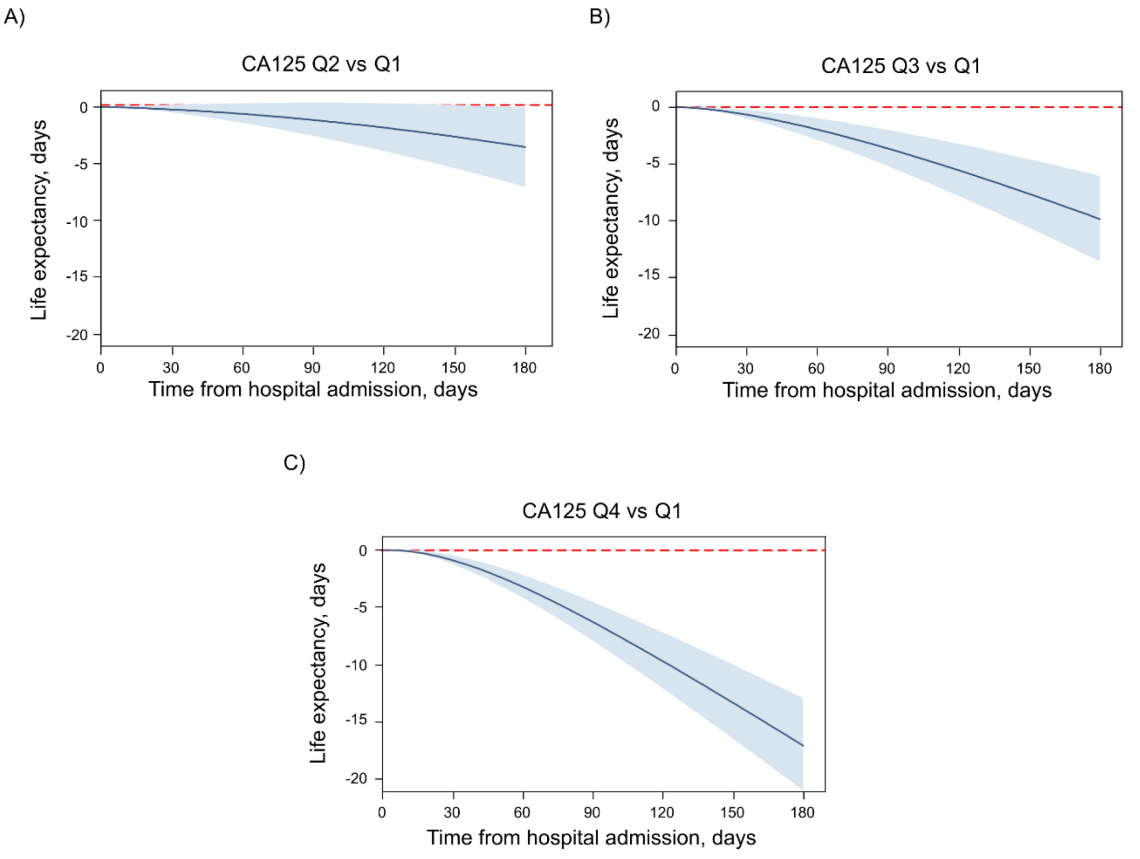
Figure 5 of the supplementary data. Time-varying hazard ratio for all-cause mortality per doubling of CA125.



The solid line represents the estimated hazard ratio for all-cause mortality per doubling of CA125 (equivalent to a 0.693-unit increase in log-transformed CA125), and the shaded band shows the corresponding 95% confidence interval. Hazard ratios are plotted on a logarithmic scale. The excess risk peaked early after discharge (reaching approximately 1.45 within the first 10 days), attenuated by day 60 (HR ≈ 1.15), and thereafter approached the null, consistent with a significant interaction with time ($P < .001$).

CA125, carbohydrate antigen 125.

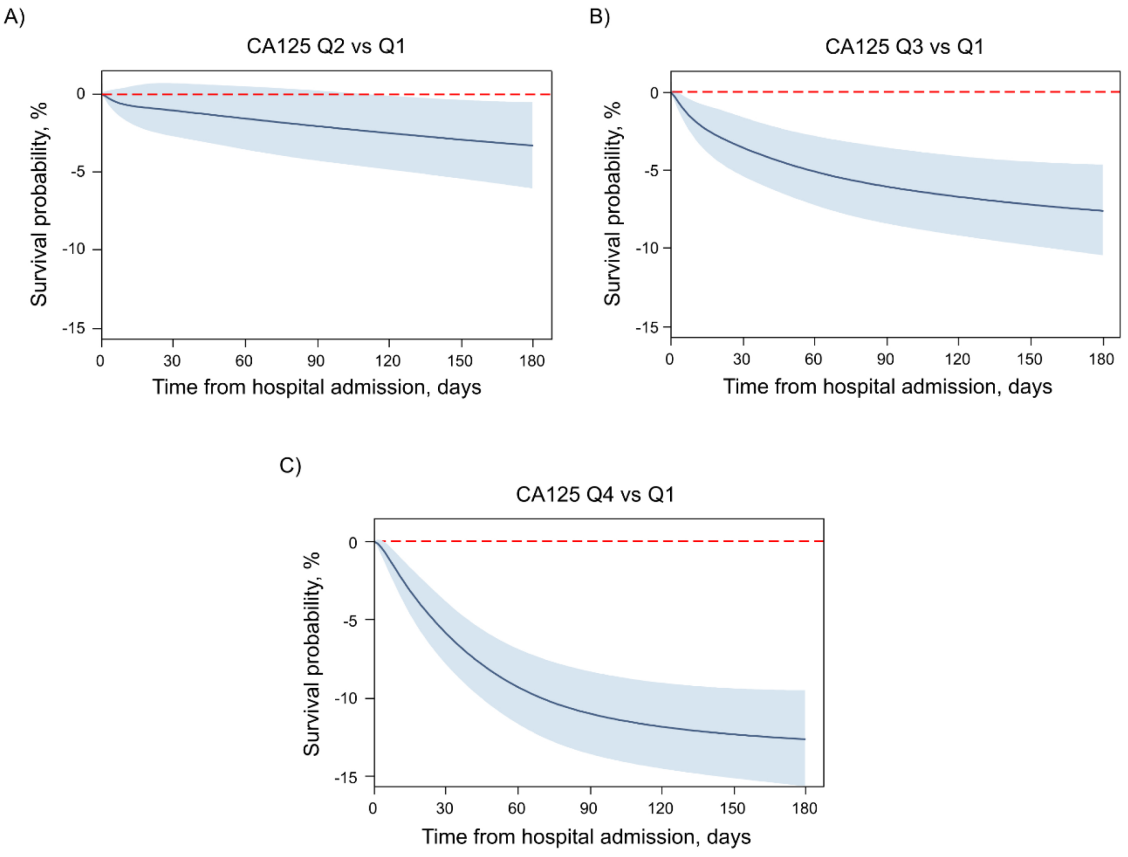
Figure 6 of the supplementary data. Differences in restricted mean survival time (RMST) up to 180 days by CA125 quartiles compared with quartile 1 (reference).



Panels display contrasts for Q2 vs Q1, Q3 vs Q1, and Q4 vs Q1. Solid lines represent estimated differences, and shaded areas indicate 95% confidence intervals. Higher CA125 quartiles were associated with progressively fewer survival days (A-C) over 180 days. Analyses were adjusted for demographic and clinical covariates. Administrative censoring occurred at 180 days.

CA125, carbohydrate antigen 125; Q, quartile.

Figure 7 of the supplementary data. Differences in standardized survival probability up to 180 days by CA125 quartiles compared with quartile 1 (reference).



Panels display contrasts for Q2 vs Q1, Q3 vs Q1, and Q4 vs Q1. Solid lines represent estimated differences, and shaded areas indicate 95% confidence intervals. Higher CA125 quartiles were associated with progressively lower survival probabilities (A-C). Analyses were adjusted for demographic and clinical covariates. Administrative censoring occurred at 180 days.

SUPPLEMENTARY REFERENCES

1. Crowther MJ. merlin—A unified modeling framework for data analysis and methods development in Stata. *Stata J Promot Commun Stat Stata*. 2020;20(4):763–784.

2. Syriopoulou E, Mozumder SI, Rutherford MJ, Lambert PC. Estimating causal effects in the presence of competing events using regression standardisation with the Stata command standsurv. *BMC Med Res Methodol*. 2022;22(1):226.