

## SUPPLEMENTARY DATA

### Supplementary methods

#### Estimation details of our multistate Markov model

We estimated our multistate Markov model via likelihood maximization based on a numerical method, using *crude initial transition frequency estimates* as an initial value of the estimation, as recommended.<sup>22</sup> We also tested the robustness of our estimated parameters by changing initial values and confirmed the robustness of our parameter estimation procedure.

#### The R code for multistate model estimation

##### (1) The multistate Markov model in figure 2

```
fitted.msm <- msm (state ~ years, subject=deid, data = df, qmatrix = Q, deathexact = c(5,6,7), censor = 999,  
censor.states = list(c(1,2,3,4)), control = list(fnscale = 5000,maxit=10000,reltol=1e-30,trace=1, REPORT=1),)
```

msm: The R function provided by the *msm* package to estimate the multistate Markov model.

#### ✕ Parameters and arguments of msm

state	The states of the multistate Markov model <ul style="list-style-type: none"><li>· 1: Excellent (KCCQ-CSS 75-100)</li><li>· 2: Good (KCCQ-CSS 50-74)</li><li>· 3: Fair (KCCQ-CSS 25-49)</li><li>· 4: Poor (KCCQ-CSS 0-24)</li><li>· 5: CV death (SCD)</li><li>· 6: CV death (non-SCD)</li><li>· 7: non-CV death</li></ul>
years	Time observed states (in year)
subject	Unique code that identifies participants

data	Dataset used
qmatrix	Initial matrix (we used the <i>crude initial transition frequency</i> estimates stated earlier)
deathexact	Exactly observed absorbing states that end the follow-up
censor	Censoring indicator: 999
censor.states	Lists the possible states which censored observations can represent.
fnscale	A scaling factor to normalize the likelihood
maxit	Maximum number of iterations
reitol	Convergence criteria
trace	Printing option to show intermediate steps of optimization in order to monitor the optimization process
REPORT	Allows following of iterations
<b>(2) The multistate Markov model in figure 3</b>	
<pre>fitted.msm &lt;- msm (state ~ years, subject=deid, covariates = ~ treatment, data = df, qmatrix = Q, deathexact = 5, censor = 999, censor.states = list(c(1,2,3,4)), control = list(fnscale = 500,maxit=10000,reitol=1e-16,trace=1,REPORT=1))</pre>	

msm: The R function provided by the *msm* package to estimate the multistate Markov model.

✂ **Parameters and arguments of msm**

state	<p>The states of the multistate Markov model</p> <ul style="list-style-type: none"> <li>· 1: Excellent (KCCQ-CSS 75-100)</li> <li>· 2: Good (KCCQ-CSS 50-74)</li> <li>· 3: Fair (KCCQ-CSS 25-49)</li> <li>· 4: Poor (KCCQ-CSS 0-24)</li> <li>· 5: All cause of death</li> </ul>
years	Time observed states (in year)
subject	Unique code that identifies participants

covariates	A formula representing the covariate
data	Dataset used
qmatrix	Initial matrix (we used <i>crude initial transition frequency</i> estimates stated earlier)
deathexact	Exactly observed absorbing states that end the follow-up:
censor	Censoring indicator: 999
censor.states	Lists the possible states which censored observations can represent.
fnscale	A scaling factor to normalise the likelihood
maxit	Maximum number of iterations
reltol	Convergence criteria
trace	Printing option to show intermediate steps of optimization in order to monitor optimization process
REPORT	Allows following of iterations

### **Markov property of the KCCQ health states**

Our multistate model assumes that the future evolution of KCCQ health states and subsequent clinical events depends only on the current KCCQ health state. This approach might seem counterintuitive to the expectation that understanding a patient's health trajectory relies on the history of health states. However, a recent study of serial KCCQ assessments showed that although prior and current KCCQ scores, as well as their differences, were all significantly associated with subsequent clinical events, when the current visit KCCQ score was included in the model together with either the prior visit score or the change from prior to current score, only the current KCCQ score remained significantly associated with prediction of subsequent clinical events.<sup>9</sup> These results indicate that the most recent KCCQ score is the most prognostically important assessment, providing evidence of the Markov property of health status dynamics in patients with heart failure with preserved ejection fraction. We further tested the fit of our Markov model in the PARAGON-HF dataset by plotting the observed and estimated proportions of individuals in each state (supplementary figure 2). The resulting plots showed that the Markov model fit the data well.

### **Specification of the transition intensity matrix (Q-matrix)**

Following best practices for multistate modeling of ordinal health states, we specified the transition intensity matrix (Q-matrix) to allow only transitions between adjacent health states. Specifically, patients could transition between adjacent states (Excellent $\leftrightarrow$ Good, Good $\leftrightarrow$ Fair, and Fair $\leftrightarrow$ Poor), while death was possible from any state. Direct transitions across non-adjacent states (eg, Excellent to Poor) were not permitted in the Q-matrix, as such observed changes in the dataset necessarily represent progression through intermediate health states that were unobserved between discrete assessment times. The likelihood calculation in the *msm* package accounts for these hidden transitions when estimating transition intensities.<sup>14,22</sup>

### **Handling of missing data and sensitivity analysis**

## #1. Primary analysis: missing-at-random (MAR) assumption

The multistate Markov model parameters were estimated using maximum likelihood estimation (MLE) via the `msm` package in R. For the primary analysis, we assumed a MAR mechanism. This approach utilizes all valid pairs of observed transitions to estimate transition intensities without requiring explicit imputation for intermittent missing values, yielding valid estimates under the MAR assumption. This assumption is supported by the study design, in which KCCQ assessments were performed at pre-specified protocol visits (baseline and at 4, 8, 12, 24, and 36 months) rather than being driven by patient symptoms.

Missing data were primarily driven by censoring events. To strictly distinguish administrative censoring or loss to follow-up from the absorbing state of death, we explicitly defined the censoring mechanism in our model by setting the `sensor.states` parameter to `list(c(1,2,3,4))`. This specification explicitly defines censored patients as being alive in one of the four health states (Excellent to Poor) at the time of censoring, allowing us to fully utilize the available survival information while mitigating potential survival bias. In the PARAGON-HF cohort, the vast majority of patients were either administratively censored (85.4%, n=4,096) or reached the absorbing state of death (14.4%, n = 691), while actual loss to follow-up with respect to vital status was negligible (< 0.2%, n = 9).

## #2. Sensitivity analysis: missing-not-at-random (MNAR) assumption

To assess the robustness of the study findings against potential informative censoring (MNAR), we performed a sensitivity analysis using a pattern-mixture model approach. While the primary analysis assumed MAR, it is possible that patients with missing assessments (including those censored at the final visit) might have experienced unobserved health deterioration.

To address this, we applied a "**worst-case deterioration**" scenario. In this stress-test analysis, all censored observations were explicitly imputed as having deteriorated to the next worse health state at the time of censoring, compared with their last observed state.

- **Imputation rule:**  $State_{censored} = State_{last\_observed} + 1$  (e.g., Good  $\rightarrow$  Fair).

- **Exception:** Patients previously in the "Poor" state (State 4) were maintained as "Poor" rather than imputed as dead, to strictly distinguish administrative censoring (alive) from the absorbing state of death.

### #3. Results of sensitivity analysis

The multistate Markov model was re-estimated using this imputed dataset. As expected, imposing artificial deterioration resulted in shorter sojourn times for better health states compared with the primary analysis. However, the comparative efficacy of sacubitril/valsartan versus valsartan remained consistent with the primary analysis.

#### Comparison of transition probabilities (primary analysis vs sensitivity analysis)

Parameter	Primary analysis (MAR)	Sensitivity analysis (MNA)	Interpretation (direction & magnitude)
<b>1. Health state dynamics (mean sojourn time)</b>			
Excellent state	1.57 y (1.51-1.63)	1.23 y (1.23-1.23)	<b>Stability decreased</b> (-0.34 y) (Confirmed rigorous stress-test)
Good state	0.70 y (0.68-0.72)	0.39 y (0.39-0.39)	<b>Stability markedly decreased</b> (-0.31 y)
Fair state	0.72 y (0.69-0.74)	0.57 y (0.57-0.57)	<b>Stability decreased</b> (-0.15 y)
Poor state	0.80 y (0.73-0.87)	1.06 y (1.05-1.06)	Artifactual Increase (+0.26 y) (Due to imputation rule: censored 'Poor' patients were retained as 'Poor', not imputed to death)
<b>2. Treatment effect</b>			
Transition: Poor → Death			
Sacubitril/valsartan	6.9% (3.3-14.1)	7.5% (3.7-14.7)	<b>Beneficial effect preserved</b>
Valsartan	13.1% (8.2-21.2)	14.1% (9.0-21.1)	
<b>Relative comparison</b>	S/V better than Val	S/V better than Val	<b>Conclusion is robust</b>

S/V, sacubitril/valsartan; Val, valsartan.

Note: The sensitivity analysis confirmed that even under the conservative assumption that all missing patients experienced health deterioration, the favorable association of sacubitril/valsartan with reduced mortality risk in patients with poor health status persisted.

#### **#4. Assessment of potential bias due to missing data**

Missing data and the timing of KCCQ assessments during follow-up pose challenges. Given that actual loss to follow-up regarding vital status was negligible (< 0.2%, n = 9) compared with the vast majority of patients who were either administratively censored (85.4%, n = 4096) or reached the absorbing state of death (14.4%, n = 691), the potential bias from informative censoring with respect to survival is likely minimal. However, regarding health status, we acknowledge the potential for informative missingness, whereby patients with worsening symptoms might be more likely to miss KCCQ assessments despite being alive. Theoretically, this could lead to an overestimation of health state stability, as reflected by longer sojourn times in the primary analysis. To address this concern, sensitivity analysis under a worst-case MNAR assumption, in which missing observations were imputed as deterioration, demonstrated that although absolute sojourn times decreased as expected, the relative favorable patterns observed with sacubitril/valsartan, particularly the lower mortality risk for patients in the poor health state, remained robust and consistent. Nevertheless, we acknowledge that statistical imputation cannot fully replace observed data, and the potential for residual bias due to unmeasured factors influencing missingness cannot be entirely eliminated

#### **Sensitivity analysis: minimal clinically important difference (MCID)-based responder and outcome analyses**

To complement the multistate Markov model and provide clinically actionable findings, we performed a landmark analysis at 8 months. This time point was selected to ensure sufficient follow-up for observing health status changes while minimizing immortal time bias.

Patients were categorized based on changes in KCCQ-CSS from baseline to 8 months using the MCID of 5 points: improvement (plus 5 points or greater), stability (change between minus 5 and plus 5), and deterioration (minus 5 points or less).

1. **Responder analysis:** the effect of sacubitril/valsartan vs valsartan on achieving a clinically meaningful improvement ( $\geq 5$  points) was assessed using a multivariable logistic regression model. The model was adjusted for baseline KCCQ-CSS, age, gender, and region.

2. **Prognostic association:** the association between KCCQ-CSS change groups at 8 months and subsequent cardiovascular death was evaluated using a Cox proportional hazards model. The model was stratified by region and adjusted for randomized treatment, age, and baseline KCCQ-CSS. Only patients alive and event-free at the 8-month landmark were included in this analysis.

#### **Potential for reverse causality**

Although our analysis adhered to the intention-to-treat principle, thereby mitigating bias arising from postrandomization changes in treatment adherence, we cannot completely exclude the possibility of reverse causality, in which worsening health status leads to reduced adherence and influences subsequent clinical outcomes.

**Table S1.** Observed transition counts between health states at consecutive visits

From\to	Excellent	Good	Fair	Poor	CV death (SCD)	CV death (Non-SCD)	Non-CV death	Censored
Excellent	10,126	2,231	308	25	67	89	83	187
Good	2,051	4,973	1,341	75	47	79	101	207
Fair	215	1,099	1,811	268	34	66	77	125
Poor	11	60	199	208	6	24	13	26

CV, cardiovascular; SCD: sudden cardiac death

Values represent the raw number of observed state pairs between consecutive scheduled visits.

Transitions to nonadjacent states (eg, excellent to fair) were handled in the continuous-time Markov model as multistep transitions occurring within the interval.

**Table S2.** Sensitivity analysis of the Markov model using KCCQ-CSS tertiles

	Tertile 3 (KCCQ-CSS: 81-100)	Tertile 2 (KCCQ-CSS: 57-80)	Tertile 1 (KCCQ-CSS: 0-56)
<b>Sojourn time, y</b>	1.21 (1.12-1.30)	0.62 (0.60-0.65)	0.95 (0.88-1.02)
<b>Next state probability, %</b>			
<i>Improved state</i>	N/A	48.8 (46.2-51.3)	90.9 (89.4-91.7)
<i>Deteriorated state</i>	96.0 (95.7-96.6)	48.5 (45.9-51.2)	N/A
<b>Probability of transitioning to death, %</b>	3.9 (3.2-4.4)	2.7 (1.6-2.9)	9.0 (8.1-11.0)
<i>Cardiovascular death</i>	2.6 (2.2-3.0)	1.7 (1.1-2.0)	5.0 (4.5-6.1)
· SCD	1.1 (1.0-1.2)	0.7 (0.5-0.7)	1.6 (1.4-1.8)
· Non-SCD	1.5 (1.2-1.7)	1.0 (0.6-1.1)	3.4 (3.1-4.1)
<i>Noncardiovascular death</i>	1.3 (1.0-1.5)	1.0 (0.5-1.1)	4.0 (3.6-5.1)
<b>Annualized death rate, %/y</b>	3.2 (2.5-4.1)	4.4 (3.2-6.0)	9.5 (8.0-11.2)

KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score N/A, non-available; SCD,

sudden cardiac death.

Values are presented with 95% confidence intervals.

**Table S3.** Proportion of patients achieving minimal clinically important differences in KCCQ-CSS at 8 months

Outcome	Sacubitril/valsartan (n = 2250)	Valsartan (n = 2226)	Overall (n = 4476)	Adjusted OR (95%CI) <sup>1</sup>	P
Improvement (≥ 5-point increase in KCCQ-CSS)	743 (33.0)	658 (29.6)	1401 (31.3)	1.17 (1.03-1.34)	.020
Deterioration (≥ 5-point decrease in KCCQ-CSS or death <sup>2</sup> )	754 (33.5)	769 (34.5)	1523 (34.0)	0.96 (0.85-1.09)	.511

CI, confidence interval; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score;

OR, odds ratio.

<sup>1</sup>Adjusted for baseline KCCQ-CSS, age, gender, and region.

<sup>2</sup>Deterioration includes patients who died before the 8-month landmark.

**Table S4.** Prognostic value of KCCQ-CSS changes at 8 months (landmark analysis)

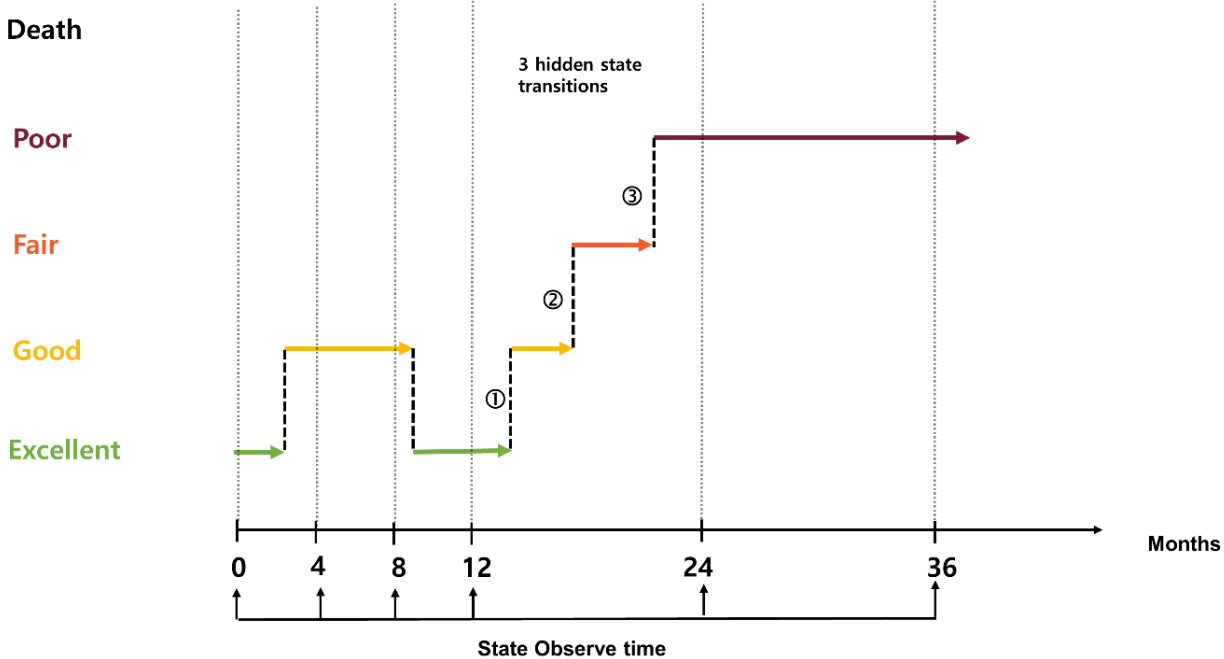
KCCQ change group (vs stable)	Adjusted hazard ratio for cardiovascular death (95% confidence interval)*	<i>P</i>	Interpretation
<b>Deteriorated</b> (≤ -5 points)	1.31 (1.00-1.71)	.048	Significantly higher risk
<b>Improved</b> (≥ +5 points)	0.76 (0.56-1.03)	.077	Trend toward lower risk

KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score.

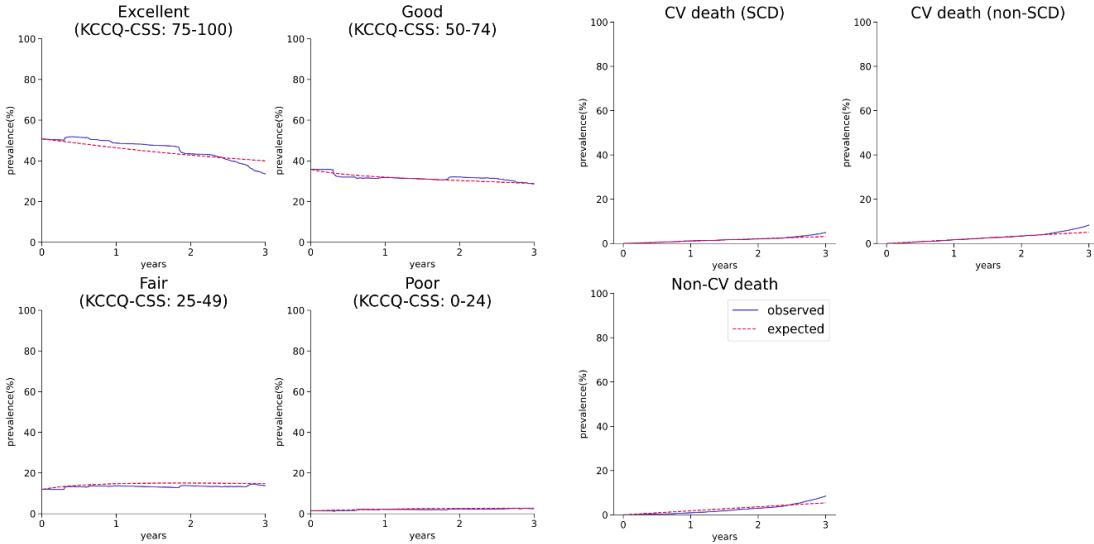
The association between minimal clinically important difference-based changes in KCCQ-CSS at 8 months and subsequent cardiovascular death was evaluated using survival analysis.

\* Adjusted for treatment, baseline KCCQ, age, and region.

Figure S1. An example of state evolution in the multistate Markov model



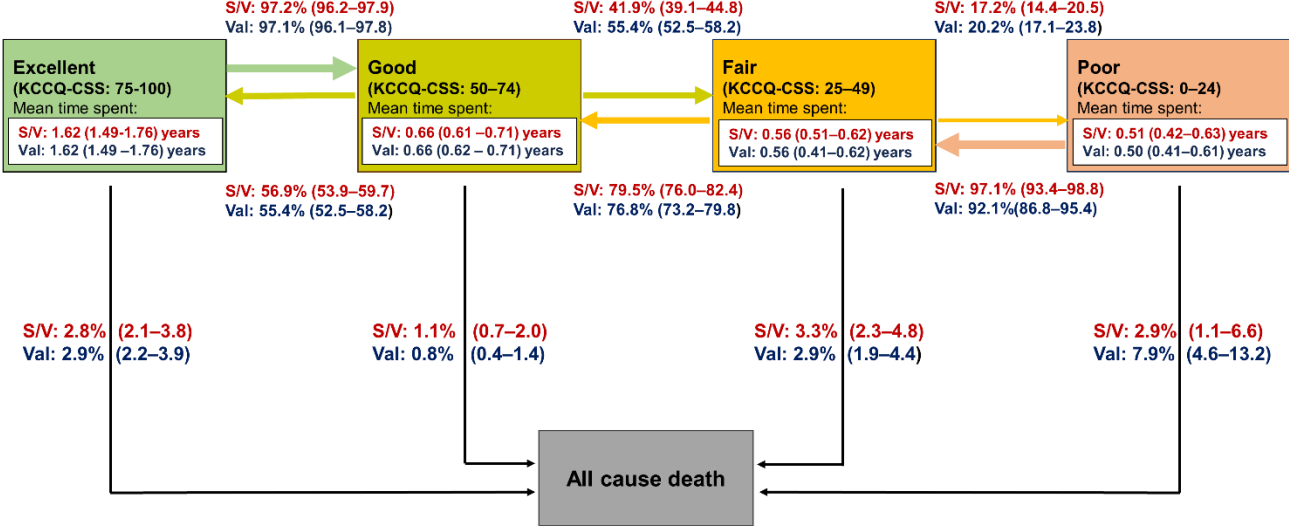
**Figure S2.** Observed and estimated prevalence of each KCCQ-CSS health state and absorbing states during follow-up



The blue line indicates the observed percentages, and the red dashed line represents the expected percentages. Time in years is shown on the x-axis, and the prevalence of each state is shown on the y-axis. The plot of observed and estimated prevalence of each KCCQ-CSS health state during follow-up suggests that our model fits the data well.

CV, cardiovascular; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score; SCD, sudden cardiac death.

**Figure S3.** Sensitivity analysis of the treatment effect of sacubitril/valsartan on health state transitions adjusted for key covariates



Age, sex, diabetes, atrial fibrillation, ischemic etiology, and prior heart failure hospitalization was adjusted in the multistate Markov model.

KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score; S/V, sacubitril/valsartan; Val, valsartan.

### Supplementary References

1. Jackson C. Multi-state models for panel data: the msm package for R. *Journal of statistical software*. 2011;38:1-28.
2. Pokharel Y, Khariton Y, Tang Y, et al. Association of Serial Kansas City Cardiomyopathy Questionnaire Assessments With Death and Hospitalization in Patients With Heart Failure With Preserved and Reduced Ejection Fraction: A Secondary Analysis of 2 Randomized Clinical Trials. *JAMA Cardiol*. 2017;2:1315-1321.
3. Jackson CH, Sharples LD, Thompson SG, Duffy SW, Couto E. Multistate Markov models for disease progression with classification error. *Journal of the Royal Statistical Society Series D: The Statistician*. 2003;52:193-209.