Aortic Stenosis and Heart Failure: Disease Ascertainment and Statistical Considerations for Clinical Trials

Ernest Spitzer,[1,2] Rebecca T Hahn,[3,4] Philippe Pibarot,[5] Ton de Vries,[2] Jeroen J Bax,[6]
Martin B Leon[3,4] and Nicolas M Van Mieghem[1]

1. Thoraxcenter, Erasmus University Medical Center, Rotterdam, the Netherlands; 2. Cardialysis, Clinical Trial Management and Core Laboratories, Rotterdam, the Netherlands; 3. New York Presbyterian Hospital/Columbia University Medical Center, New York, NY, US; 4. Cardiovascular Research Foundation, New York, NY, US; 5. Quebec Heart and Lung Institute, Laval University, Quebec, Canada; 6. Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands

Abstract

Aortic stenosis is a progressive disease that develops over decades, and once symptomatic and untreated, is associated with poor survival. Transcatheter aortic valve replacement has evolved significantly in the past decade and has expanded its indication from surgically inoperable and high-risk patients to patients with intermediate risk. Assessment of heart failure-related outcomes include the use of functional assessments, disease-specific quality of life surveys and standardised ascertainment of events, such as hospitalisations. Multiple statistical approaches are currently being tested to account for recurrent events such as hospitalisations for heart failure or to combine binary and continuous outcomes, both intended to assess the holistic burden of the disease, as opposed to the traditional analysis of time to first event.
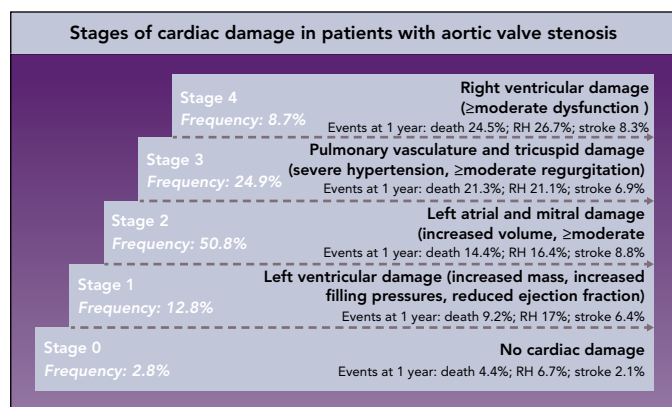
Symptomatic severe aortic stenosis (AS) is the most common indication for valvular interventions.[1] AS is a degenerative and progressive disease that characteristically remains asymptomatic for decades but once symptoms occur, survival is severely compromised. Historical data have shown that the time from the onset of symptoms to death is about 2 years in patients who develop heart failure (HF) symptoms, 3 years in those who present with a syncope and 5 years in those presenting with angina.[2] The Long-term follow-up of the Placement of Aortic Transcatheter Valves (PARTNER 1B) trial showed that two-thirds of inoperable patients who followed standard treatment did not survive beyond 2 years, while transcatheter aortic valve replacement (TAVR) halved mortality.[3]

Stages of cardiac damage in patients with severe AS have recently been defined (*Figure 1*).[4] Stage 1 includes increased left ventricular mass, increased left ventricular filling pressures and systolic dysfunction defined as left ventricular ejection fraction (LVEF) <50%. Further stages relate to damage of the left atrium or mitral valve (stage 2), pulmonary vasculature or tricuspid valve (stage 3) and right ventricular damage (stage 4). Each stage is associated with an increased risk of mortality within 1 year, ranging from 4% at stage 0 (no damage) up to 25% at stage 4. Stage 1 patients, when compared with AS patients

without cardiac damage, show an increased mortality (9% versus 4%), hospitalisation rate (17% versus 7%) and stroke rate (6% versus 2%). Recent studies have called into question the traditional 50% LVEF cut-off, suggesting that in patients with AS, an ejection fraction of ≤60% may precede the onset of symptoms and may also predict progression of the disease.[5] Thus, evaluation of the left ventricular systolic function is critical in the follow-up of patients with asymptomatic AS and early detection of dysfunction prompts the need for accelerated aortic valve replacement.[1] Moreover, ascertainment of history of congestive HF may better characterise the prognosis of patients who are having a planned intervention. The impact of chronic HF and systolic dysfunction on treatment selection – TAVR versus surgical aortic valve replacement (SAVR) – merits further research.[1]

In this review, we summarise the prevalence of HF in patients included in clinical trials comparing TAVR with SAVR or medical treatment, as well as tools for assessment of the functional status, quality of life and clinical events during follow-up. We also discuss recent recommendations for broadening the definition of HF-related clinical events as well as statistical methods that not only increase the power of comparisons, but also may better capture the burden of the disease and effects of experimental therapies in the setting of clinical trials.

Co-morbidities

Figure 1: Stages of Cardiac Damage In Patients With Aortic Valve Stenosis

Stages of cardiac damage in patients with aortic valve stenosis

Stage 4
Frequency: 8.7%
Right ventricular damage (≥moderate dysfunction)
Events at 1 year: death 24.5%; RH 26.7%; stroke 8.3%

Stage 3
Frequency: 24.9%
Pulmonary vasculature and tricuspid damage (severe hypertension, ≥moderate regurgitation)
Events at 1 year: death 21.3%; RH 21.1%; stroke 6.9%

Stage 2
Frequency: 50.8%
Left atrial and mitral damage (increased volume, ≥moderate
Events at 1 year: death 14.4%; RH 16.4%; stroke 8.8%

Stage 1
Frequency: 12.8%
Left ventricular damage (increased mass, increased filling pressures, reduced ejection fraction)
Events at 1 year: death 9.2%; RH 17%; stroke 6.4%

Stage 0
Frequency: 2.8%
No cardiac damage
Events at 1 year: death 4.4%; RH 6.7%; stroke 2.1%

Frequency of each stage and observed rates of events at 1 year are based on data from the PARTNER 2A and 2B trials.[22,23] RH = rehospitalisation.

Cardiac imaging to assess LVEF, longitudinal strain, mitral regurgitation and hemodynamic parameters, such as pulmonary artery systolic pressure, as well as biomarkers, such as brain natriuretic peptide (BNP) and N-terminal pro-BNP (NT-proBNP), have been associated with HF symptoms and worsening symptoms, but detailed discussion of these is beyond the scope of this review.[6]

Prevalence of Heart Failure and Related Comorbidities

HF is multifactorial in patients with severe AS and can be a consequence of the increased afterload and myocardial remodeling, with the contributory effect of cardiac damage characteristics of stages 2–4, or secondary to ischaemia.[4] Characterising the aetiology requires interrogation and assessment of previous MI or coronary artery disease, status of coronary artery lesions (i.e. existence of lesions requiring intervention), atrial fibrillation and pulmonary hypertension. Defining prior congestive HF is not standardised and may range from ambulatory symptoms prior to hospitalisations for HF. In *Table 1* we summarise the baseline characteristics of seven clinical trials, providing data for both the TAVR and control groups when available.[7–13] Prior MI ranged from 5% in a study with an all-comers design to 31% in an extreme risk cohort;[14] coronary artery disease affected two-thirds of patients and atrial fibrillation one-third.[7–13] Remarkably, previous HF was captured only in three of the seven studies, and was highly prevalent (≥95%) in patients with intermediate, high or extreme risk.[7–12] LVEF was reported in four of seven studies and mean values were always above the cut-off value accepted for normality (>50%). An ejection fraction <50% was seen in ~30–50% of patients with severe AS.[7–9,12] The New York Heart Association (NYHA) functional class at baseline was used in all studies and reflected accurately the risk of the analysed cohorts. In an all-comers design,[13] approximately half of the patients presented with NYHA class I or II (*Table 1*), while this number was less than 10% in cohorts with high or extreme risk.[7,8,12] Likewise, NYHA class IV was present in up to half of patients at high surgical risk, while it was observed in <3% in the all-comers cohort.[12,13] These findings underscore the value of NYHA class for characterising the baseline functional status of patients with severe AS. Although the reproducibility of this assessment has been criticised, its simplicity and availability make it a useful functional assessment.[15] It is noteworthy that prior congestive HF should be better defined and standardised and consistently captured in cardiovascular trials.

Ascertainment of Heart Failure-related Clinical Events at Follow-up
All-cause and Cardiovascular Mortality

TAVR has revolutionised the management of severe AS. This is largely due to continued improvement in transcatheter heart valves and implantation techniques. Efforts to expand its indication have targeted populations with progressively lower surgical risk.[7–13] These combined factors resulted in a consistent decrease in overall rates of all-cause death at 1 year from 31% (n=179) in the inoperable cohort of the PARTNER IB trial treated with TAVR,[8] to 7% (n=864) in the Surgical Replacement and Transcatheter Aortic Valve Implantation (SURTAVI) trial targeting patients with an intermediate risk, and 5% (n=145) in the all-comers Nordic Aortic Valve Intervention (NOTION) trial (*Table 2*).[11,13] In cardiovascular research, all-cause mortality is considered the most robust and unbiased clinical endpoint (*Figure 2*).[16] Nevertheless, it may lack specificity, and thus differentiation between cardiovascular and non-cardiovascular death is compulsory. Given the complexity in classifying events as cardiovascular or non-cardiovascular, the involvement of an independent clinical events committee is considered a quality marker when interpreting trial outcomes.[16] It has been suggested that using cardiovascular death in composite primary endpoints instead of all-cause mortality, for example cardiovascular death and hospitalisations for HF, reduces statistical noise generated by non-cardiovascular fatal events that are generally not influenced by targeted cardiovascular interventions.[17]

Hospitalisation due to Heart Failure

Although rehospitalisation due to HF is considered a less robust endpoint in clinical trials due to the lack of implementation of standardised definitions, it remains the most important outcome from patient prognosis and health economic perspectives. In three of the seven trials included in this review it was not reported, and definitions slightly varied when it was available.[7,10,13] Frequently, it is difficult to distinguish between a hospitalisation due to aortic valve disease and/or complications of the valve procedure versus a hospitalisation due to HF. These are not always mutually exclusive and strict criteria should be applied to be able to adjudicate and report both. A standardised definition of hospitalisation due to HF is needed if meaningful comparison of rates among cardiovascular trials are to be made.[16] The Standardised Data Collection for Cardiovascular Trials Initiative (SCTI), in collaboration with the Food and Drug Administration (FDA), the American College of Cardiology and the American Heart Association recommend a standardised definition for HF events, which include urgent, unscheduled outpatient office/practice, emergency department visits and hospitalisations due to HF.[18] HF hospitalisation occurs when a patient is admitted to the hospital with a primary diagnosis of HF, the length of stay is at least 24 hours (or extends over a calendar date), the patient exhibits at least one new or worsening symptom of HF, has objective evidence of new or worsening HF (at least two signs or one sign and one laboratory finding), and receives initiation or intensification of treatment specifically for HF. HF signs and symptoms, relevant changes in therapy, as well as laboratory findings – BNP or NT-proBNP, radiological evidence, non-invasive cardiac imaging, right heart catheterisation – are carefully defined in the SCTI document. The almost simultaneous publication of the Mitral Valve Academic Research Consortium consensus manuscript defines what qualifies as a hospitalisation (≥24 hour stay) with criteria for HF hospitalisation or rehospitalisation requiring signs, symptoms and/or laboratory evidence of worsening HF and administration of IV or mechanical HF therapies.[19] They further sub-classify HF hospitalisation into primary (cardiac-related) and secondary (non-cardiac related).

Table 1: Baseline Characteristics in Patients Included in Clinical Trials Investigating Transcatheter Aortic Valve Replacement in Severe Aortic Stenosis

Clinical Trial	PARTNER IB (Inoperable)[8]		CoreValve (Extreme Risk)[10,14]	PARTNER IA (High Risk)[12]		CoreValve (High Risk)[7,14]		PARTNER IIA (Intermediate Risk)[9]		SURTAVI (Intermediate Risk)[11]		NOTION (All-comers)[13]	
	TAVR	Standard therapy	TAVR	TAVR	SAVR	TAVR	SAVR	TAVR	SAVR	TAVR	SAVR	TAVR	SAVR
Patients analysed	179	179	487	348	351	394	401	1,011	1,021	864	796	145	135
Age in years	83	83	83	84	85	83	84	82	82	80	80	79	79
Women (%)	54	53	52	42	43	46	47	46	45	42	45	46	47
Prior MI (%)	19	26	31	27	30	26	24	18	18	15	14	6	4
Prior PCI (%)	31	25	37	34	33	34	38	27	28	21	21	8	9
Prior CABG (%)	37	46	40	43	44	30	30	24	26	16	17	NA	NA
Coronary artery disease (%)	68	74	82	75	77	75	76	69	67	63	64	NA	NA
AF (%)	33	49	47	41	43	41	48	31	35	28	27	28	28
Significant mitral regurgitation (%)	22	23	NA	20	21	NA	NA	17	19	NA	NA	NA	NA
Pulmonary hypertension (%)	42	44	NA	42	36	NA	NA	NA	NA	NA	NA	NA	NA
Congestive heart failure (%)	NA	NA	97	NA	NA	95	97	NA	NA	95	97	NA	NA
LVEF (%)	54 ±13	51 ±14	55 ±14	53 ±14	53 ±13	NA	NA	56 ±11	55 ±12	NA	NA	NA	NA
LVEF <50% (%)	38	47	38	43	40	NA	NA	28	33	NA	NA	NA	NA
NYHA class*													
I (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.9	2.2
II (%)	7.7	5.8	8.2	5.6	5.4	14.2	13.2	22.2	24.1	39.8	41.8	46.5	52.2
III (%)	48.7	48.6	64.0	40.9	42.8	65.5	69.1	60.2	57.4	54.6	51.7	46.5	42.6
IV (%)	43.6	45.6	27.8	53.4	51.8	20.3	17.7	17.6	18.5	5.6	6.5	2.1	3.0

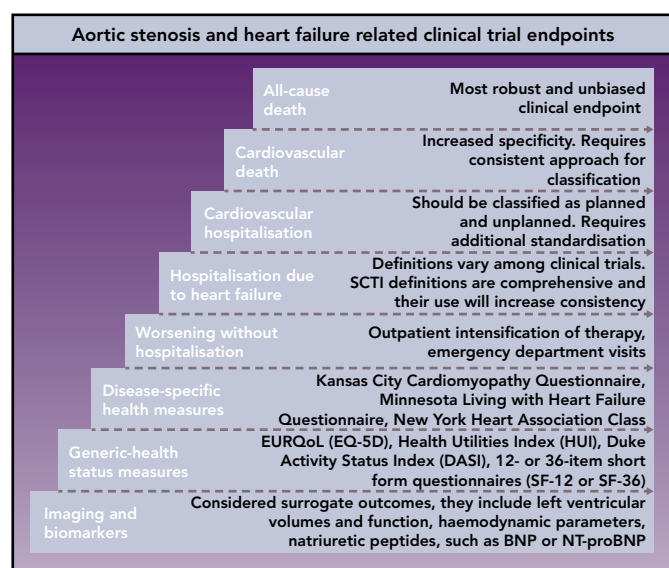*Numbers were derived from frequency figures when numerical data were not available.
CABG = coronary artery bypass grafting; LVEF = left ventricular ejection fraction; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; TAVR = transcatheter aortic valve replacement; SAVR = surgical aortic valve replacement.

Table 2: Heart Failure-Related Events up to 1 Year of Follow-up in Clinical Trials Investigating Transcatheter Aortic Valve Replacement in Severe Aortic Stenosis

	PARTNER IB (Inoperable)[8]		CoreValve (Extreme Risk)[10,14]	PARTNER IA (High Risk)[12]		CoreValve (High Risk)[7,14]		PARTNER IIA (Intermediate Risk)[9]		SURTAVI (Intermediate Risk)[11]		NOTION (All-comers)[13]	
	TAVR	Standard therapy	TAVR	TAVR	SAVR	TAVR	SAVR	TAVR	SAVR	TAVR	SAVR	TAVR	SAVR
Patients analysed	179	179	487	348	351	394	401	1011	1021	864	796	145	135
All-cause death	30.7	49.7	24.3	24.2	26.8	13.9	18.7	12.3	12.9	6.7	6.8	4.9	7.5
Cardiovascular death	19.6	41.9	18.3	14.3	13	NA	NA	7.1*	8.1*	4.8	5.5	4.3	7.5
Rehospitalisation	22.3	44.1	NA	18.2	15.5	NA	NA	14.8	14.7	8.5	7.6	NA	NA
Change in % LVEF	0.0	−12.0	2.8	4.0	3.4	NA	NA	−0.3	2.1	NA	NA	NA	NA
NYHA class change†													
I (%)	24.25	1.13	42.75	36.12	33.88	48.00	44.00	55.09	58.33	70.70	68.20	62.40	79.40
II (%)	17.08	11.13	15.22	21.98	24.31	14.00	10.00	3.24	−2.31	−15.80	−14.90	−18.00	−37.60
III (%)	−34.34	−29.72	−58.70	−30.69	−35.78	−65.00	−61.00	−54.63	−52.31	−50.20	−49.10	−43.10	−38.80
IV (%)	−41.23	−38.96	−25.72	−51.98	−50.00	−13.00	−14.00	−16.67	−17.59	−4.50	−4.20	−1.40	−3.00
Death or missing at 1 year† (%)	34.25	56.42	26.45	24.57	27.59	16.00	22.00	12.96	13.89	7.00	6.80	4.90	7.50
KCCQ change at 1 year	NA	NA	27 (24–31)‡	29 (24–33)‡§	27 (22–32)‡	23 ±26	22 ±27	22 (20–24)‡§	22 (20–24)‡	21 ±22	21 ±22	NA	NA

*Only cardiac causes were included. †Numbers were derived from frequency figures when numerical data were not available. ‡Mean value (95% CI). §Numbers reflect the transfemoral cohort. KCCQ = Kansas City Cardiomyopathy Questionnaire; LVEF = left ventricular ejection fraction; NYHA = New York Heart Association; SAVR = surgical aortic valve replacement; TAVR = transcatheter aortic valve replacement.

Figure 2: Hierarchy of Clinical Trial Endpoints Related to Aortic Stenosis and Heart Failure

Aortic stenosis and heart failure related clinical trial endpoints	
All-cause death	Most robust and unbiased clinical endpoint
Cardiovascular death	Increased specificity. Requires consistent approach for classification
Cardiovascular hospitalisation	Should be classified as planned and unplanned. Requires additional standardisation
Hospitalisation due to heart failure	Definitions vary among clinical trials. SCTI definitions are comprehensive and their use will increase consistency
Worsening without hospitalisation	Outpatient intensification of therapy, emergency department visits
Disease-specific health measures	Kansas City Cardiomyopathy Questionnaire, Minnesota Living with Heart Failure Questionnaire, New York Heart Association Class
Generic-health status measures	EURQoL (EQ-5D), Health Utilities Index (HUI), Duke Activity Status Index (DASI), 12- or 36-item short form questionnaires (SF-12 or SF-36)
Imaging and biomarkers	Considered surrogate outcomes, they include left ventricular volumes and function, haemodynamic parameters, natriuretic peptides, such as BNP or NT-proBNP

CEC = clinical events committee; SCTI = Standardized data Collection for Clinical Trials Initiative.

A recent sub-analysis of the Prospective Comparison of ARNI with ACEI to Determine the Impact of Global Mortality and Morbidity in Heart Failure (PARADIGM-HF) trial, a randomised, double-blind comparison of sacubitril/valsartan with enalapril in 8,399 patients with chronic HF, showed that patients hospitalised due to HF had a significantly increased risk of all-cause death (HR 5.0; 95% CI [4.4–5.7]) throughout the duration of the trial (27 months) in an adjusted analysis for randomised treatment, region and baseline covariates, when analysing hospitalisation for HF as the only event experienced as a time-updated covariate.[20] When this analysis was carried out for emergency department visits due to HF (without subsequent hospitalisation), the risk of all-cause death was three times higher than in patients without an event (HR 2.9; 95% CI [1.9–4.6]). The intensification of HF therapy as an outpatient was also evaluated, since many episodes of worsening HF are treated in the community with an increase in oral pharmacological therapy or the use of short-term IV therapy.[20] When analysing this endpoint as the only event experienced as a time-updated covariate, the authors observed an increased risk of death (HR 4.2; 95% CI [3.3–5.3]), almost equivalent to that observed in patients hospitalised due to HF. These findings did not only clarify the prognostic role of HF events not linked to hospitalisation, but further showed that adding intensification of HF therapy and ED visits due to HF, the frequency of HF-related events doubled, suggesting that an extended composite endpoint would increase statistical power without compromising specificity, which is especially appealing for event-driven clinical trials. These data further support the implementation of SCTI-defined HF events.[18]

Health Status Measures
The impact of transcatheter therapies for severe AS on functional capacity has been largely assessed by changes in NYHA class and, less frequently, by the use of validated disease-specific questionnaires such as the Kansas City Cardiomyopathy Questionnaire (KCCQ).[7–13] When interpreting comparisons of cross-sectional measures, it is important to take into account that generally subjects will be not all be present for any specific follow-up time point, due to death,

a missed appointment or patients lost to follow-up. Consequently, these comparisons will unequivocally exclude patients who are the most ill. Aortic valve replacement has significantly and consistently increased the number of patients classified as NYHA class I, ranging from a 24% increase relative to baseline in inoperable patients, to an 80% increase in all-comer populations at one-year follow-up.[8,13] Likewise, the frequency of NYHA class III and IV has been reduced up to 65% and 50% respectively in high-risk cohorts (Table 2).[7,12]

The KCCQ is a 23-item, self-administered instrument that quantifies physical function, symptoms (frequency, severity and recent changes), social function, self-efficacy and knowledge and quality of life. The instrument provides a score from 0 to 100. It has shown an excellent correlation with NYHA class and each quartile is reflective of an increased risk of mortality in patients with HF. It has been validated for the assessment of prognosis and effects of therapies in severe AS.[21] A recent sub-analysis of the PARTNER 2 trial evaluated changes in KCCQ at 1 month, 1 year and 2 years among patients at intermediate risk randomised to TAVR or SAVR.[22] For this analysis the authors categorised changes in KCCQ as follows: death, worse (reduction from baseline >5 points), no change (change between –5 and <5 points), mildly improved (increase between 5 and <10 points), moderately improved (increase between 10 and <20 points), and substantially improved (increase ≥20 points). Overall, there was a similar increase in KCCQ at 1 year in the transfemoral TAVR group (22.1 points; 20.4–23.9) and in the SAVR group (22.1 points; 20.1–24.1), albeit an earlier benefit was observed in patients undergoing transfemoral TAVR. Moreover, the frequency of moderate or substantial improvement (≥10 points in KCCQ) was consistent among groups (71.1% in the TAVR group and 68% in the SAVR group). Similar findings have been reported in the trials included in this review.[11,14,21–23] Other instruments of proven value for the assessment of health status are the disease-specific Minnesota Living with Heart Failure Questionnaire and generic health status measures such as the EuroQoL, Health Utilities Index, Duke Activity Status Index, 12- or 36-item short-form questionnaires (SF-12 or SF-36).[15,24,25]

Statistical Analysis of Heart Failure Endpoints
Clinical primary endpoints in HF and AS trials are customarily analysed using a Kaplan–Meier analysis using the log-rank test and the treatment effect calculated with the Cox proportional hazards regression using one of several tests, such as the Wald test. These methods are well established and used in studies with regulatory approval studies.[26] Primary endpoints in HF trials include, for example, all-cause death; all-cause death and hospitalisations for HF; and more recently cardiovascular death and hospitalisations for HF. Likewise, AS trials have used all-cause death or all-cause death and stroke. The time-to-first-event analysis allows the most intuitive presentation of results, but does not fully capture the burden of recurrent events, such as hospitalisations for HF, an issue that becomes more relevant when investigating patients with AS and impaired systolic function, as envisioned in the ongoing Transcatheter Aortic Valve Replacement to UNload the Left Ventricle in Patients With ADvanced Heart Failure (TAVR UNLOAD) trial.[27] The high frequency of recurrent HF events is denoted in trials such as the PARADIGM HF, in which one-third of patients who were hospitalised once during follow-up were hospitalised at least for a second time throughout the duration of the trial, and one tenth were hospitalised three or more times.[28] Similar distribution of recurrent events have been reported in drug and device intervention trials.[29–33] Patients with more unplanned

Table 3: Statistical Methods for the Analysis of Recurrent Events used in Heart Failure Research

Statistical Methods for the Analysis of Recurrent Events Used in Heart Failure Research			
Time-to-event models These methods are extensions of the Cox proportional hazards model; preferred when censoring exists; all models except the join frailty model consider all events as a single-event process; most are supported by standard statistical software.	**Joint frailty model (JFM) (1998)** Based on total time scale; semi-parametric or parametric; accounts for the dependence between recurrent events and death through a patient-specific frailty term; able to 1:1 model time-varying covariates; assumes a constant event rate and treatment effect over time (e.g. NCT01626079)	**Count models** Compare counts of events in a given time; not preferred when the event rate is not constant over time or when there are time-varying effects	**Poisson regression (1837)*** Assumes that the underlying event rate is the same in all subjects and that event counts follow the Poisson distribution (the mean and variance are equal); recurrent events are assumed to be independent; does not incorporate competing risk.
	Lin, Wei, Ying and Yang (LWYY) model (2000) Based on gap time; stratified Cox-based; recurrent events are assumed to be independent; semi-parametric; terminal events are assumed to be recurrent events; does not incorporate competing risks or time-varying covariates into the analysis (e.g. NCT01920711)		**Negative binomial regression (1714)*** Assumes an association between recurrent events through a random effect term; does not incorporate competing risks; assumes a constant event rate and treatment effect over time; allows to sdjust the variance independently from the mean (e.g. NCT00531661)
	Prentice–Williams–Peterson (PWP) model (1981) Based on gap times or total time scale; stratified Cox-based; semi-parametric; terminal events are assumed to be recurrent events; incorporates the order of the events; conditional model; recurrent events risk can be influenced by previous events	**Cumulative incidence methods**	**Gosh and Lin method (2001)** Terminal events are handled as informative censoring (competing risk); non-parametric and semi-parametric; models the mean frequency function **Nelson–Aelen estimator (1969)** Terminal events are handled as independent censoring
	Andersen-Gill (AG) model (1982) Based on gap times; semi-paramedic; recurrent events are assumed to be independent; terminal events are assumed to be recurrent events; does not incorporate: competing risks; can include time-varying covariates; robust variance (e.g. NCT0053166)	**Generalised pairwise comparison methods** Non-parametric ranked tests applied on composite endpoints allowing continuous, ordinal and censored events	**O'Brien method (non-hierarchical) (1984)** Compares multiple outcomes based on an overall rank for each subject and using a rank-sum or ANOVA test **Finkelstein–Schoenfeld method (hierarchical) (1999)** Compares binary and continuous outcomes in a hierarchy; may include functional assessments or imaging 1 parameters (e.g. NCT02661451; NCT01768702) **Buyse method (hierarchical) (2010)** Extension of the Wilcoxon–Mann–Whitney test **Matched/unmatched Pocock method (Win ratio) (2012)** For every pair, a 'winner' or a 'loser' is defined. The win ratio is defined by winners/losers (e.g. NCT00530894)
	Wei, Lin and Weissfe/d (WLW) model (1989) Based on total time scale; stratified Cox-based; semi-parametric; recurrent events are assumed to be independent; unconditional marginal model; each event has its own stratum; ignores the order of the events; does not incorporate competing risks		

Year of introduction of the distributions; all other years refer for first report for each method/model.

hospitalisations exhibit a worse quality of life and survival, thus, being able to analyse recurrent hospitalisations not only better characterises the disease but also increases statistical power to detect differences in treatment effects.

Since the early 1980s, several statistical approaches to account for multiple hospitalisations have been introduced (*Table 3*), and recently a shift towards these more complex methods has been observed in trials enrolling patients with HF, aiming for efficiency and robustness.[28,34–36] For instance, in the groundbreaking Cardiovascular Outcomes Assessment of the MitraClip Percutaneous Therapy for Heart Failure Patients with Functional Mitral Regurgitation (COAPT) trial, the chosen primary effectiveness endpoint was all hospitalisations for HF within 24 months of follow-up, including recurrent events in patients with more than one event, using the joint frailty model to account for correlated events and the competing risk of death.[37] This is one of the available time-to-event approaches, which include the Wei, Lin and Weissfeld (WLW) method; the Lin, Wei, Ying and Yang (LWYY) model; the Prentice–Williams–Peterson model; and the Andersen–Gill

model.[28,38] The choice for the most appropriate statistical approach relates to:

- The distribution of timing of subsequent events – HF rehospitalisations may not occur after similar intervals but in clusters where some patients will present multiple adjacent episodes and others no recurrences.
- The within-patient correlation of subsequent events – it is known that hospitalisations beget more hospitalisations and worse prognosis, thus methods assuming independence of recurrent events may not be preferred for analysis of HF events.
- Frequency of the recurrent and terminal events – where methods that analyse death as non-informative censoring or as a recurrent event may not be ideal for cohorts in which mortality is expected to be relatively high.

A conservative approach is to include a method for a primary statistical analysis based on the study assumptions, and provide sensitivity analyses based on other methods for robustness.

Co-morbidities

Alternative approaches include the use of methods based on event rates, such as the Poisson regression and negative binomial regression. The latter allows for more flexibility but assuming a constant event rate over time and not analysing death as a competing risk, thus it may not be preferred in scenarios where fatal events account for a high proportion of the composite. Of interest in this situation is the Gosh and Lin cumulative incidence method, which handles fatal events as informative censoring.[39] In a recent pre-specified sub-analysis of the PARADIGM-HF trial, the authors compared results of the analysis of recurrent hospitalisations using a cumulative incidence method, time-to-event models (WLW, LWYY and the joint frailty model) and the negative binomial model. All approaches provided similar estimates for the effect of the experimental therapy (sacubitril/valsartan) when compared with the traditional time-to-first-event analysis (log-rank test).[28] The authors concluded that no single method can be recommended over another, and the preferred statistical approach for a specific trial should be discussed with regulatory agencies.[26] It is noteworthy that the joint frailty and the LWYY methods offer advantages that have prompted their use in recent studies (*Table 3*).[17,28,40]

Generalised pairwise comparison (GPC) methods have been developed, which use non-parametric approaches to compare outcomes on the basis of pairs of subjects.[41] Hierarchical GPC methods include the Finkelstein–Schoenfeld method, the unmatched Pocock method (or win ratio) and the Buyse method, among others.[35,42,43]

These methods allow the creation of a hierarchy that gives a higher priority to the most severe outcomes and are able to accommodate multiple events. Characteristically, GPC methods are used for binary outcomes, such as in the primary endpoint of the Tafamidis in Transthyretin Cardiomyopathy Clinical Trial (ATTR-ACT) trial, which included a hierarchical assessment of all-cause death and frequency of cardiovascular-related hospitalisations or the combination of binary and continuous outcomes, such as in the primary endpoint of the TAVR UNLOAD trial, defined as the hierarchical occurrence within one year of all-cause death, disabling stroke, hospitalisations (related to HF, symptomatic aortic valve disease or non-disabling stroke) and change in KCCQ relative to baseline.[27,44] Non-hierarchical GPC methods include the O'Brien method.[45] Little is known about the relative benefits of one method over another. The Finkelstein–Schoenfeld method is currently the GPC method most widely used in cardiovascular research.

Conclusion

HF events, impaired functional status and reduced disease-specific quality of life are highly prevalent in patients with aortic stenosis and are significantly and positively affected by aortic valve interventions. The use of standardised definitions for HF-related events is recommended to improve our understanding of the disease and to allow comparisons among clinical trials. Further research on complex statistical approaches, which take into account the occurrence of multiple events, is warranted. ■

1. Baumgartner H, Falk V, Bax JJ, et al. 2017 ESC/EACTS Guidelines for the management of valvular heart disease. *Eur Heart J* 2017;38:2739–91. https://doi.org/10.1093/eurheartj/ehx391; PMID: 28886619.
2. Frank S, Johnson A, Ross J Jr. Natural history of valvular aortic stenosis. *Br Heart J* 1973;35:41–6. https://doi.org/10.1136/hrt.35.1.41; PMID: 4685905.
3. Kapadia SR, Leon MB, Makkar RR, et al. 5-year outcomes of transcatheter aortic valve replacement compared with standard treatment for patients with inoperable aortic stenosis (PARTNER 1): a randomised controlled trial. *Lancet* 2015;385:2485–91. https://doi.org/10.1016/S0140-6736(15)60290-2; PMID: 25788231.
4. Genereux P, Pibarot P, Redfors B, et al. Staging classification of aortic stenosis based on the extent of cardiac damage. *Eur Heart J* 2017;38:3351–8. https://doi.org/10.1093/eurheartj/ehx381; PMID: 29020232.
5. Ito S, Miranda WR, Nkomo VT, et al. Reduced left ventricular ejection fraction in patients with aortic stenosis. *J Am Coll Cardiol* 2018;71:1313–21. https://doi.org/10.1016/j.jacc.2018.01.045; PMID: 29566814.
6. Everett RJ, Clavel MA, Pibarot P, Dweck MR. Timing of intervention in aortic stenosis: a review of current and future strategies. *Heart* 2018;104:2067–76 https://doi.org/10.1136/heartjnl-2017-312304; PMID: 30030337.
7. Adams DH, Popma JJ, Reardon MJ, et al. Transcatheter aortic-valve replacement with a self-expanding prosthesis. *N Engl J Med* 2014;370:1790–8. https://doi.org/10.1056/NEJMoa1400590; PMID: 24678937.
8. Leon MB, Smith CR, Mack M, et al. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. *N Engl J Med* 2010;363:1597–607. https://doi.org/10.1056/NEJMoa1008232; PMID: 20961243.
9. Leon MB, Smith CR, Mack MJ, et al. Transcatheter or surgical aortic-valve replacement in intermediate-risk patients. *N Engl J Med* 2016;374:1609–20. https://doi.org/10.1056/NEJMoa1514616; PMID: 27040324.
10. Popma JJ, Adams DH, Reardon MJ, et al. Transcatheter aortic valve replacement using a self-expanding bioprosthesis in patients with severe aortic stenosis at extreme risk for surgery. *J Am Coll Cardiol* 2014;63:1972–81. https://doi.org/10.1016/j.jacc.2014.02.556; PMID: 24657695.
11. Reardon MJ, Van Mieghem NM, Popma JJ, et al. Surgical or transcatheter aortic-valve replacement in intermediate-risk patients. *N Engl J Med* 2017;376:1321–31. https://doi.org/10.1056/NEJMoa1700456; PMID: 28304219.
12. Smith CR, Leon MB, Mack MJ, et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. *N Engl J Med* 2011;364:2187–98. https://doi.org/10.1056/NEJMoa1103510; PMID: 21639811.
13. Thyregod HG, Steinbruchel DA, Ihlemann N, et al. Transcatheter versus surgical aortic valve replacement in patients with severe aortic valve stenosis: 1-year results

from the all-comers NOTION randomized clinical trial. *J Am Coll Cardiol* 2015;65:2184–94. https://doi.org/10.1016/j.jacc.2015.03.014; PMID: 25787196.
14. Osnabrugge RL, Arnold SV, Reynolds MR, et al. Health status after transcatheter aortic valve replacement in patients at extreme surgical risk: results from the CoreValve US trial. *JACC Cardiovasc Interv* 2015;8:315–23. https://doi.org/10.1016/j.jcin.2014.08.016; PMID: 25700755.
15. Allen LA, Spertus JA. End points for comparative effectiveness research in heart failure. *Heart Fail Clin* 2013;9:15–28. https://doi.org/10.1016/j.hfc.2012.09.002; PMID: 23168314.
16. Zannad F, Garcia AA, Anker SD, et al. Clinical outcome endpoints in heart failure trials: a European Society of Cardiology Heart Failure Association consensus document. *Eur J Heart Fail* 2013;15:1082–94. https://doi.org/10.1093/eurjhf/hft095; PMID: 23787718.
17. Shen L, Jhund PS, Mogensen UM, et al. Re-examination of the BEST trial using composite outcomes, including emergency department visits. *JACC Heart Fail* 2017;5:591–9. https://doi.org/10.1016/j.jchf.2017.04.005; PMID: 28774394.
18. Hicks KA, Tcheng JE, Bozkurt B, et al. 2014 ACC/AHA key data elements and definitions for cardiovascular endpoint events in clinical trials: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (writing committee to develop cardiovascular endpoints data standards). *J Am Coll Cardiol* 2015;66:403–69. https://doi.org/10.1016/j.jacc.2014.12.018; PMID: 25553722.
19. Stone GW, Adams DH, Abraham WT, et al. Clinical trial design principles and endpoint definitions for transcatheter mitral valve repair and replacement: part 2: endpoint definitions: a consensus document from the Mitral Valve Academic Research Consortium. *J Am Coll Cardiol* 2015;66:308–21. https://doi.org/10.1016/j.jacc.2015.05.049; PMID: 26184623.
20. Okumura N, Jhund PS, Gong J, et al. Importance of clinical worsening of heart failure treated in the outpatient setting: evidence from the prospective comparison of ARNI with ACEI to determine impact on global mortality and morbidity in heart failure trial (PARADIGM-HF). *Circulation* 2016;133:2254–62. https://doi.org/10.1161/CIRCULATIONAHA.115.020729; PMID: 27143684.
21. Arnold SV, Spertus JA, Lei Y, et al. Use of the Kansas City Cardiomyopathy Questionnaire for monitoring health status in patients with aortic stenosis. *Circ Heart Fail* 2013;6:61–7. https://doi.org/10.1161/CIRCHEARTFAILURE.112.970053; PMID: 23230306.
22. Baron SJ, Arnold SV, Wang K, et al. Health status benefits of transcatheter vs surgical aortic valve replacement in patients with severe aortic stenosis at intermediate surgical risk: results from the PARTNER 2 randomized clinical trial. *JAMA Cardiol* 2017;2:837–45. https://doi.org/10.1001/jamacardio.2017.2039; PMID: 28658491.
23. Reynolds MR, Magnuson EA, Wang K, et al. Health-related quality of life after transcatheter or surgical aortic valve

replacement in high-risk patients with severe aortic stenosis: results from the PARTNER (Placement of AoRTic TraNscathetER Valve) Trial (Cohort A). *J Am Coll Cardiol* 2012;60:548–58. https://doi.org/10.1016/j.jacc.2012.03.075; PMID: 22818074.
24. Supino PG, Borer JS, Franciosa JA, et al. Acceptability and psychometric properties of the Minnesota Living With Heart Failure Questionnaire among patients undergoing heart valve surgery: validation and comparison with SF-36. *J Card Fail* 2009;15:267–77. https://doi.org/10.1016/j.cardfail.2008.10.003; PMID: 19327629.
25. Kelkar AA, Spertus J, Pang P, et al. Utility of patient-reported outcome instruments in heart failure. *JACC Heart Fail* 2016;4:165–75. https://doi.org/10.1016/j.jchf.2015.10.015; PMID: 26874386.
26. European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of chronic heart failure. Available at: http://www.ema.europa.eu/ema/pages/includes/document/open_document.jsp?webContentId=WC500235089 (accessed 15 November 2017).
27. Spitzer E, Van Mieghem NM, Pibarot P, et al. Rationale and design of the Transcatheter Aortic Valve Replacement to UNload the Left ventricle in patients with ADvanced heart failure (TAVR UNLOAD) trial. *Am Heart J* 2016;182:80–8. https://doi.org/10.1016/j.ahj.2016.08.009; PMID: 27914503.
28. Mogensen UM, Gong J, Jhund PS, et al. Effect of sacubitril/valsartan on recurrent events in the Prospective comparison of ARNI with ACEI to Determine Impact on Global Mortality and morbidity in Heart Failure trial (PARADIGM-HF). *Eur J Heart Fail* 2018;20:760–8. https://doi.org/10.1002/ejhf.1139; PMID: 29431251.
29. Rogers JK, McMurray JJ, Pocock SJ, et al. Eplerenone in patients with systolic heart failure and mild symptoms: analysis of repeat hospitalizations. *Circulation* 2012;126:2317–23. https://doi.org/10.1161/CIRCULATIONAHA.112.110536; PMID: 23042980.
30. Rogers JK, Pocock SJ, McMurray JJ, et al. Analysing recurrent hospitalizations in heart failure: a review of statistical methodology, with application to CHARM-Preserved. *Eur J Heart Fail* 2014;16:33–40. https://doi.org/10.1002/ejhf.29; PMID: 24453096.
31. Abraham WT, Adamson PB, Bourge RC, et al. Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial. *Lancet* 2011;377:658–66. https://doi.org/10.1016/S0140-6736(11)60101-3; PMID: 21315441.
32. Claggett B, Pocock S, Wei LJ, et al. Comparison of time-to-first event and recurrent-event methods in randomized clinical trials. *Circulation* 2018;138:570–7. https://doi.org/10.1161/CIRCULATIONAHA.117.033065; PMID: 29588314.
33. Pitt B, Pfeffer MA, Assmann SF, et al. Spironolactone for heart failure with preserved ejection fraction. *N Engl J Med*

2014;370:1383–92. https://doi.org/10.1056/NEJMoa1313731; PMID: 24716680.

34. Ozga AK, Kieser M, Rauch G. A systematic comparison of recurrent event models for application to composite endpoints. *BMC Med Res Methodol* 2018;18:2. https://doi.org/10.1186/s12874-017-0462-x; PMID: 29301487.

35. Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med* 1999;18:1341–54. https://doi.org/10.1002/(SICI)1097-0258(19990615)18:11<1341::AID-SIM129>3.0.CO;2-7; MID: 10399200.

36. Ullah S, Gabbett TJ, Finch CF. Statistical modelling for recurrent events: an application to sports injuries. *Br J Sports Med* 2014;48:1287–93. https://doi.org/10.1136/bjsports-2011-090803; PMID: 22872683.

37. Stone GW, Lindenfeld J, Abraham WT, et al. Transcatheter mitral-valve repair in patients with heart failure. *N Engl J Med*

2018;379:2307–18. https://doi.org/10.1056/NEJMoa1806640; PMID: 30280640.

38. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. *Int J Epidemiol* 2015;44:324–33. https://doi.org/10.1093/ije/dyu222; PMID: 25501468.

39. Ghosh D, Lin DY. Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* 2003;59:877–85. https://doi.org/10.1111/j.0006-341X.2003.00102.x; PMID: 14969466.

40. Rogers JK, Yaroshinsky A, Pocock SJ, et al. Analysis of recurrent events with an associated informative dropout time: Application of the joint frailty model. *Stat Med* 2016;35:2195–205. https://doi.org/10.1002/sim.6853; PMID: 26751714.

41. Ramchandani R, Schoenfeld DA, Finkelstein DM. Global rank tests for multiple, possibly censored, outcomes. *Biometrics*

2016;72:926–35. https://doi.org/10.1111/biom.12475; PMID: 26812695.

42. Buyse M. Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. *Stat Med* 2010;29:3245–57. https://doi.org/10.1002/sim.3923; PMID: 21170918.

43. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J* 2012;33:176–82. https://doi.org/10.1093/eurheartj/ehr352; PMID: 21900289.

44. Maurer MS, Schwartz JH, Gundapaneni B, et al. Tafamidis treatment for patients with transthyretin amyloid cardiomyopathy. *N Engl J Med* 2018;379:1007–16. https://doi.org/10.1056/NEJMoa1805689; PMID: 30145929.

45. O'Brien PC. Procedures for comparing samples with multiple endpoints. *Biometrics* 1984;40:1079–87. https://doi.org/10.2307/2531158; PMID: 6534410.