Current perspectives of using **PED/PROs as primary end points** for regulatory approval of drugs

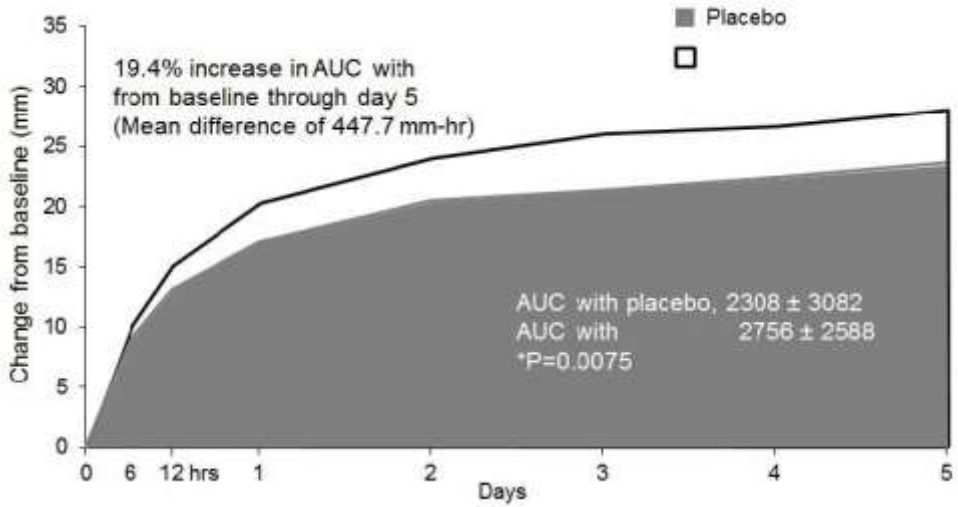
Unmet Needs in Cardiovascular Diseases focusing on patient-benefit risk and patient-reported outcomes

Patrick Vrijlandt

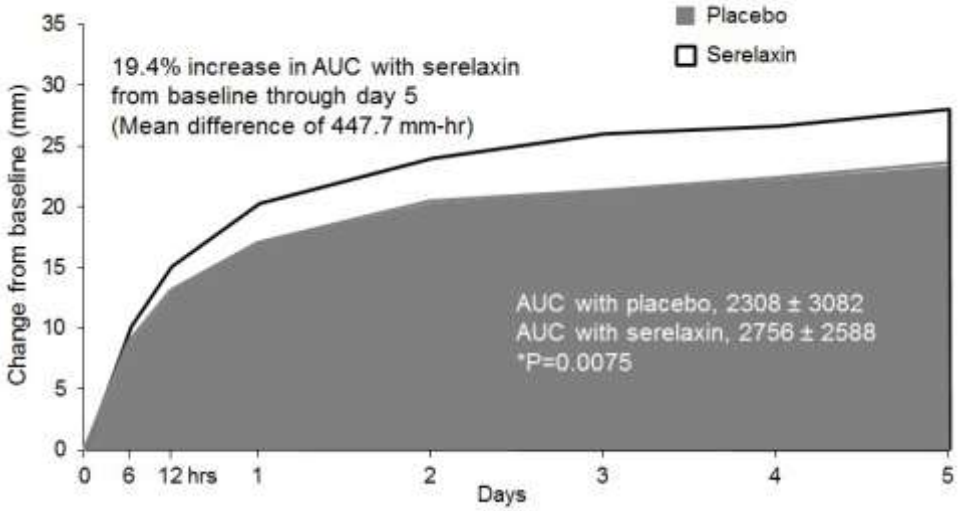
19 February 2025

ESC
European Society of Cardiology

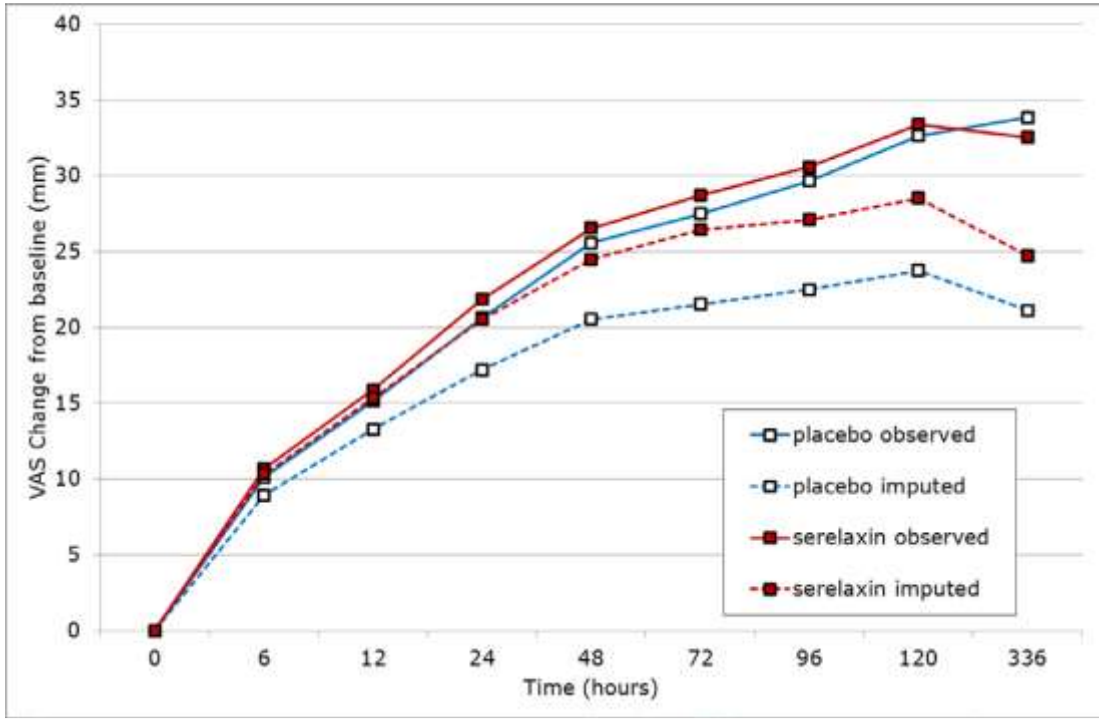
Mean change from baseline (mm) of **dyspnoea Visual Analogue Scale (VAS)**



Mean change from baseline (mm) of **dyspnoea Visual Analogue Scale (VAS) –** Study RELAX-AHF (ITT analysis set)



VAS in RELAX-AHF showing observed and imputed data



DAPA-HF → SmPC

ESC Round Table

Table 10. Number and percent of patients with clinically meaningful improvement and deterioration on the KCCQ-TSS at 8 months

Change from baseline at 8 months:	Dapagliflozin 10 mg n[a]=2086	Placebo n[a]=2062		
Improvement	**n (%) improved[b]**	**n (%) improved[b]**	**Odds ratio[c] (95% CI)**	**p-value[f]**
≥ 5 points	933 (44.7)	794 (38.5)	1.14 (1.06, 1.22)	0.0002
≥ 10 points	689 (33.0)	579 (28.1)	1.13 (1.05, 1.22)	0.0018
≥ 15 points	474 (22.7)	406 (19.7)	1.10 (1.01, 1.19)	0.0300
Deterioration	**n (%) deteriorated[d]**	**n (%) deteriorated[d]**	**Odds ratio[e] (95% CI)**	**p-value[f]**
≥ 5 points	537 (25.7)	693 (33.6)	0.84 (0.78, 0.89)	<0.0001
≥ 10 points	395 (18.9)	506 (24.5)	0.85 (0.79, 0.92)	<0.0001

[a] Number of patients with an observed KCCQ-TSS or who died prior to 8 months.

[b] Number of patients who had an observed improvement of at least 5, 10 or 15 points from baseline. Patients who died prior to the given timepoint are counted as not improved.

[c] For improvement, an odds ratio > 1 favours dapagliflozin 10 mg.

[d] Number of patients who had an observed deterioration of at least 5 or 10 points from baseline. Patients who died prior to the given timepoint are counted as deteriorated.

[e] For deterioration, an odds ratio < 1 favours dapagliflozin 10 mg.

[f] p-values are nominal.

PARADIGM-HF – sacubritil/valsartan - KCCQ-CSS

Improvement of HF symptoms and physical limitations as assessed by KCCQ

Patients in the LCZ696 group showed less reduction compared to **enalapril** from baseline to Month 8 in the clinical summary score for HF symptoms and physical limitations. The between-group mean difference for the clinical summary score was 1.64, with a 95% CI of 0.63 to 2.65 (one-sided p = 0.0007). This reduction in the decline of the clinical summary scores for LCZ696 vs enalapril did not meet the threshold for significance using the strict MTP at an alpha = 0.001 as pre-specified in the statistical analysis plan (ie, required $p \leq 0.0002$), but it met the threshold for significance using the alternative MTP (requiring one-sided $p \leq 0.00458$).

Mean change LCZ: -2.99 Enal:-4.63

EXPLORER-HCM mavacamten → SmPC

Table 4: Analysis of the primary composite and secondary endpoints from EXPLORER-HCM study

	Mavacamten N = 123	Placebo N = 128
Patients achieving primary endpoint at week 30, n (%)	45 (37%)	22 (17%)
Treatment difference (95% CI)	19.4 (8.67, 30.13)	
p-value	0.0005	
Change from baseline post-exercise LVOT peak gradient at week 30, mmHg	N = 123	N = 128
Mean (SD)	-47 (40)	-10 (30)
Treatment difference* (95% CI)	-35 (-43, -28)	
p-value	< 0.0001	
Change from baseline to week 30 in pVO_2, mL/kg/min	N = 123	N = 128
Mean (SD)	1.4 (3)	-0.05 (3)
Treatment difference* (95% CI)	1.4 (0.6, 2)	
p-value	< 0.0006	
Patients with improvement of NYHA class ≥ 1 at week 30	N = 123	N = 128
N, (%)	80 (65%)	40 (31%)
Treatment difference (95% CI)	34 (22, 45)	
p-value	< 0.0001	
Change from baseline to week 30 in KCCQ-23 CSS†	N = 92	N = 88
Mean (SD)	14 (14)	4 (14)
Treatment difference* (95% CI)	9 (5, 13)	
p-value	< 0.0001	
Baseline	N = 99	N = 97
Mean (SD)	71 (16)	71 (19)
Change from baseline to week 30 in HCMSQ SoB domain score‡	N = 85	N = 86
Mean (SD)	-2.8 (2.7)	-0.9 (2.4)
Treatment difference* (95% CI)	-1.8 (-2.4, -1.2)	
p-value	< 0.0001	
Baseline	N = 108	N = 109
Mean (SD)	4.9 (2.5)	4.5 (3.2)

* Least-squares mean difference

† KCCQ-23 CSS = Kansas City Cardiomyopathy Questionnaire-23 Clinical Summary Score. The KCCQ-23 CSS is derived from the Total Symptoms Score (TSS) and the Physical Limitations (PL) score of the KCCQ-23. The CSS ranges from 0 to 100, with higher scores representing better health status. A significant treatment effect on the KCCQ-23 CSS favouring mavacamten was first observed at week 6 and remained consistent through week 30.

‡ HCMSQ SoB = Hypertrophic Cardiomyopathy Symptom Questionnaire Shortness of Breath. The HCMSQ SoB domain score measures frequency and severity of shortness of breath. The HCMSQ SoB domain score ranges from 0 to 18, with lower scores representing less shortness of breath. A significant treatment effect on the HCMSQ SoB favouring mavacamten was first observed at week 4 and remained consistent through week 30.

Tafamidis → SmPC

Table 4: 6MWT and KCCQ-OS and component domain scores

Endpoints	Baseline Mean (SD)		Change from Baseline to Month 30, LS mean (SE)		Treatment difference from placebo LS mean (95% CI)	*p-value*
	Pooled Tafamidis N=264	Placebo N=177	Pooled Tafamidis	Placebo		
6MWT* (metres)	350.55 (121.30)	353.26 (125.98)	-54.87 (5.07)	-130.55 (9.80)	75.68 (57.56, 93.80)	$p< 0.0001$
KCCQ-OS*	67.27 (21.36)	65.90 (21.74)	-7.16 (1.42)	-20.81 (1.97)	13.65 (9.48, 17.83)	$p< 0.0001$

* Higher values indicate better health status.
Abbreviations: 6MWT=6-Minute Walk Test; KCCQ-OS=Kansas City Cardiomyopathy Questionnaire-Overall Summary; LS=least squares; CI=confidence interval.

A DUE – macitentan/tadalafil

- The difference in **6MWD** change from baseline to EDBT was not statistically significant between M/T FDC and macitentan or between M/T FDC and tadalafil. Based on the ANCOVA model run at each stage with treatment group, stratum, and baseline value as covariates, the median unbiased estimates of change from baseline to EDBT (adjusted RCL) and combined p-values were:
  - M/T FDC versus macitentan: 16.04 m (-17.0,49.08), p=0.380
  - M/T FDC versus tadalafil: 25.37 m (-0.93,51.59), p=0.059
- There were no differences in change from baseline to EDBT (treatment effect, 95% CI) between groups in the PAH-SYMPACT **cardiopulmonary** domain score:
  - -0.03 (-0.21,0.15) for M/T FDC versus macitentan
  - -0.04 (-0.21,0.13) for M/T FDC versus tadalafil
- There were no differences in change from baseline to EDBT (treatment effect, 95% CI) between groups in the PAH-SYMPACT **cardiovascular** domain score:
  - 0.01 (-0.17,0.19) for M/T FDC versus macitentan
  - 0.02 (-0.15,0.19) for M/T FDC versus tadalafil

STEP-HFpEF(-DM) semaglutide 2.4 mg → SmPC

ESC

d Table

Table 11 Results of 6MWD, KCCQ-CSS and body weight from the two 52-week randomised trials (STEP-HFpEF and STEP-HFpEF-DM)

	STEP-HFpEF		STEP-HFpEF-DM	
-	Semaglutide 2.4 mg	Placebo	Semaglutide 2.4 mg	Placebo
Full analysis set (N)	263	266	310	306
KCCQ-CSS (score)				
Baseline (mean)[1]	57.9	55.5	58.8	56.4
Change from baseline[2]	16.6	8.7	13.7	6.4
Difference from placebo[2] [95% CI]	7.8 [4.8; 10.9]		7.3 [4.1; 10.4]	
Patients (%) experiencing meaningful change[3]	43.2	32.5	42.7	30.5
6MWD (metres)				
Baseline (mean)[1]	319.6	314.6	279.7	276.7
Change from baseline[2]	21.5	1.2	12.7	-1.6
Difference from placebo[2] [95% CI]	20.3 [8.6; 32.1]		14.3 [3.7; 24.9]	
Patients (%) with meaningful change[4]	47.9	34.7	43.8	30.6
Body weight				
Baseline (kg)[1]	108.3	108.4	106.4	105.2
Change (%) from baseline[2]	-13.3	-2.6	-9.8	-3.4
Difference (%) from placebo[2] [95% CI]	-10.7 [-11.9; -9.4]		-6.4 [-7.6; -5.2]	

[1] Observed mean.
[2] Estimated using an ANCOVA model using multiple imputation and for KCCQ and 6MWD, also a composite imputation based on all data irrespective of discontinuation of randomised treatment or initiation of other anti-obesity medication or bariatric surgery.
[3] Meaningful within patient change threshold of 17.2 points for STEP-HFpEF trial and 16.3 points for STEP-HFpEF-DM trial (derived using an anchor-based method based on a 1-category improvement in Patient Global Impression of Status (PGI-S)). Percentages are based on subjects with an observation at the visit.
[4] Meaningful within patient change threshold of 22.1 metres for STEP-HFpEF trial and 25.6 metres for STEP-HFpEF-DM trial (derived using an anchor-based method using "moderately better" in Patient Global Impression of Change (PGI-C)). Percentages are based on subjects with an observation at the visit.

Conclusions / thoughts

- **Many different scales/endpoints (including subscales)**
  → we're not getting familiar
- **No generally agreed threshold for relevance or even threshold elicited in the trial**
  → we can't value the outcome
- **Many different outcome metrics**
  - Mean group difference > MCID
  - Responders > MCID (improvement, deterioration)
  - Winratio for highest change etc
- **Statistical issues**
  → missing data are not random
- **Active control**
- **Too low in hierarchy**
- **HTA**