Testimony to the FDA Cardiovascular and Renal Drugs Advisory Committee

Serelaxin

(BLA 125468)

March 27, 2014

Sammy Almashat, MD, MPH
Michael Carome, MD
Sidney Wolfe, MD
Public Citizen Health Research Group
(We have no financial conflicts of interest)

Reasons for opposing approval

- Serelaxin has exceedingly marginal, if any, benefits on subjective, transient, patient-reported dyspnea; no proven benefits on any objective cardiovascular outcomes
- Proposed indication bears little resemblance to primary endpoints studied
- Extremely poor generalizability of RELAX-AHF (~20% of all acute heart failure (AHF) inpatients would have met inclusion criteria)

Primary endpoints

- RELAX-AHF had two pre-specified co-primary endpoints, both subjective, patient-reported measures of dyspnea in the acute setting:
  - Responder analysis: Proportion of respondents with "moderately" to "markedly" improved dyspnea on a 7-point Likert scale within 24 hours after drug administration
  - VAS-AUC: Mean area under the curve (AUC) of subjective dyspnea on a 100-mm visual analog scale (VAS) over 5-day period after drug administration

FDA's standard for approval of serelaxin based on a single Phase III trial (expressed during a Feb. 3, 2009 pre-submission meeting)

"Sponsor said there would be two phase 3 studies. FDA told the sponsor that only 1 study would be required for the dyspnea indication if rejection of the global null hypothesis with the alpha level was controlled at the two-sided 0.00125 level by the Hochberg method. For this, at least one of the two dyspnea coprimary endpoints would have to have a p<0.00125 and the other endpoint would have to demonstrate a trend consistent with the results on the first endpoint." (FDA Briefing Document, p. 22)

All citations to page numbers are to the FDA Briefing Document posted in advance of the meeting, unless otherwise noted: (http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/CardiovascularandRenalDrugsAdvisoryCommittee/UCM390443.pdf)

Primary endpoint #1: Responder analysis

- Proportion of subjects with "moderate" or "marked" improvement in subjective dyspnea at 6, 12, and 24 hours after drug administration (ITT):
  - 26.9% and 25.9% in the serelaxin and placebo groups, respectively
  - OR 1.05 (95% CI: 0.81, 1.37; p=0.7)

Primary endpoint #2: VAS-AUC

- VAS-AUC at Day 5:
  - 2756 vs. 2308 mm-hrs in serelaxin vs. placebo, respectively
  - Mean difference = 448 mm-hrs
  - p=0.007 (>0.00125: FDA standard for approval without confirmatory trial)
- This mean difference was based not on *actual* VAS values for all subjects, but was instead driven *primarily* by the sponsor's chosen imputation method for subjects with worsening heart failure

Arbitrary imputation protocol for VAS scores in worsening heart failure

- All subjects with "worsening heart failure" (for which no criteria were provided *a priori* and, which was not subsequently adjudicated), regardless of severity, were automatically given worst possible VAS value (0) for all subsequent values through Day 5
- Twice as many placebo subjects had "worsening heart failure" as serelaxin subjects.
- Therefore, replacing raw VAS values with "0" for the remainder of their assessments disproportionately – and markedly – decreased VAS-AUC scores in placebo subjects relative to serelaxin subjects

Arbitrary imputation

- Two-thirds (66%) of all serelaxin and placebo subjects with "worsening heart failure" through Day 14 had mild cases, which were adequately treated with additional infusions of IV diuretics or nitrates
- FDA, after conducting 6 different sensitivity analyses employing alternative, and less extreme, imputation protocols, concluded that:
  - *"It is notable that only the [sponsor's chosen] prespecified imputation scheme which treats all degrees of severity of [worsening heart failure] equally keeps the p value below the prespecified 0.025 mark needed for success." (p. 82)*

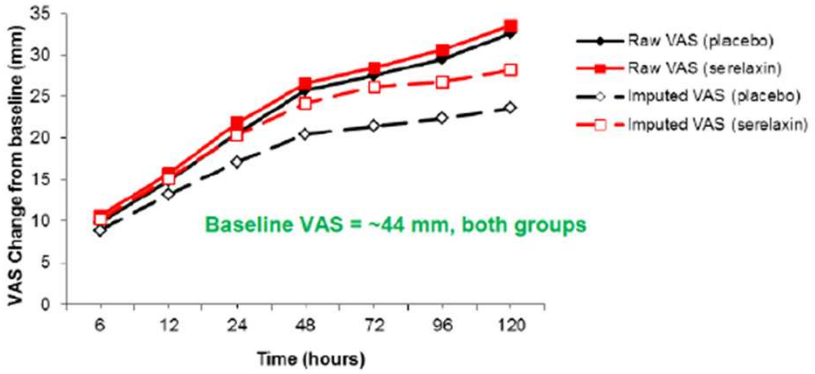
FDA Briefing Doc, p. 78. 104 of 157 (66%) subjects with "worsening heart failure" through Day 14 were treated with nothing more than IV diuretics or IV nitrates. Subjects were only counted once, with outcomes considered by decreasing severity (death, followed by ultrafiltration, mechanical ventilation, , IV milrinone, IV pressors, IV nitrates, IV diuretics, and other and finally, rehospitalization for HF).

Arbitrary imputation

- *"As shown in the FDA statistical review, if raw VAS scores are used the difference in AUC-change from baseline in VAS scores over 5 days between groups was only 167.6.mm-hr (**p = 0.21**)." (p. 67)*
- Therefore, the actual difference between the serelaxin and placebo groups was just **1.4 mm** on a 100 mm scale at any one time for the first 5 days after drug administration, far below any conceivable threshold for clinical significance.

Actual (raw) VAS differences

Figure 8: Raw (using LOCF) VAS scores and imputed VAS scores at all assessments from 6 hours through Day 5



Dr. McDowell's analysis: Source: Sponsor's analysis dataset for VAS scores

(p. 68)

Clinical significance of VAS score

- Even taking the sponsor's analysis at face value, what does the imputed mean difference in AUC-VAS really mean?
- "448 mm-hrs" over 5 days = **3.7 mm on a 100 mm scale** at any one time (448/120 hrs over 5 days).

Clinical (in)significance

Based on a 2004 study of VAS in AHF inpatients (Ander et al.) cited by FDA:

- "Patients who experience[d] **about the same** difficulty in breathing experience[d] a mean change of **2.7 mm** on the VAS (95% CI -4.2 mm to 9.5 mm)."
- "Patients who experience[d] **a little less** difficulty breathing or a little more difficulty breathing experienced a mean change of **21.1 mm** on the VAS (95% CI: 12.3 mm to 29.9 mm)." (p. 67-68)

Ander DS, Aisiku IP, Ratcliff JJ, Todd KH, Gotsch K. Measuring the dyspnea of decompensated heart failure with a visual analog scale: how much improvement is meaningful? Congest Heart Fail. 2004 Jul-Aug;10(4):188-91.
"The mean change in VAS score associated with a **minimum clinically significant change** in breathing difficulty was **21.1 mm** (95% CI, 12.3 mm–29.9 mm). The mean change in VAS score for those that reported "about the same difficulty breathing" was 2.7 mm (95% CI, –4.2 mm –9.5 mm)."

Clinical (in)significance

Dr. Blank: *"Judging by [the Ander et al.] experience in what appeared to be a similar population, a 4 mm difference on a 100 mm dyspnea VAS does not appear to represent a clinically significant change." (p. 16)*

We agree.

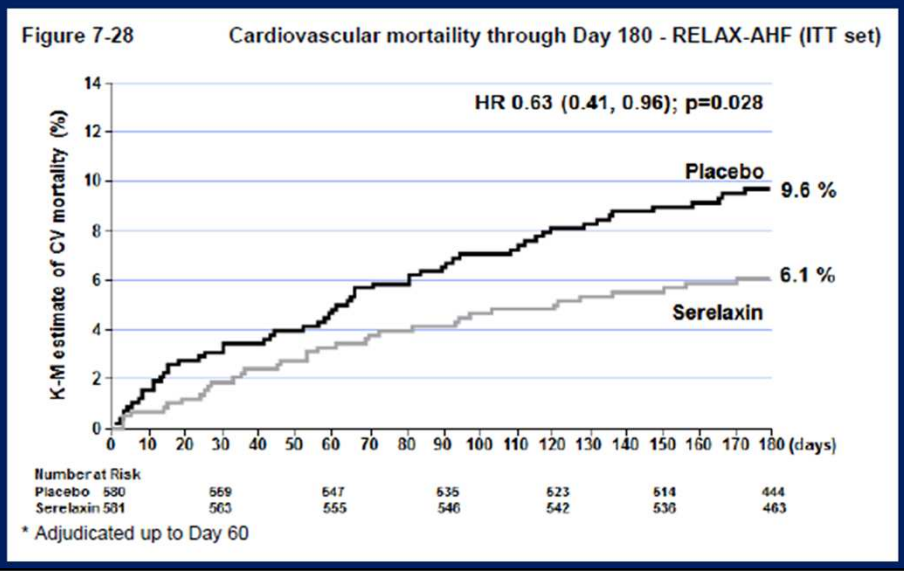
Secondary, clinical endpoints not met

- Days alive and out of hospital through Day 60
  - 48.3 days vs. 47.7 days (serelaxin vs. placebo; p=0.3682)
- Cardiovascular death or re-hospitalization due to HF or RF through Day 60
  - 13.2% vs. 13.0% (serelaxin vs. placebo); p=0.8945
  - HR: 1.02 (95% CI: 0.74, 1.41)

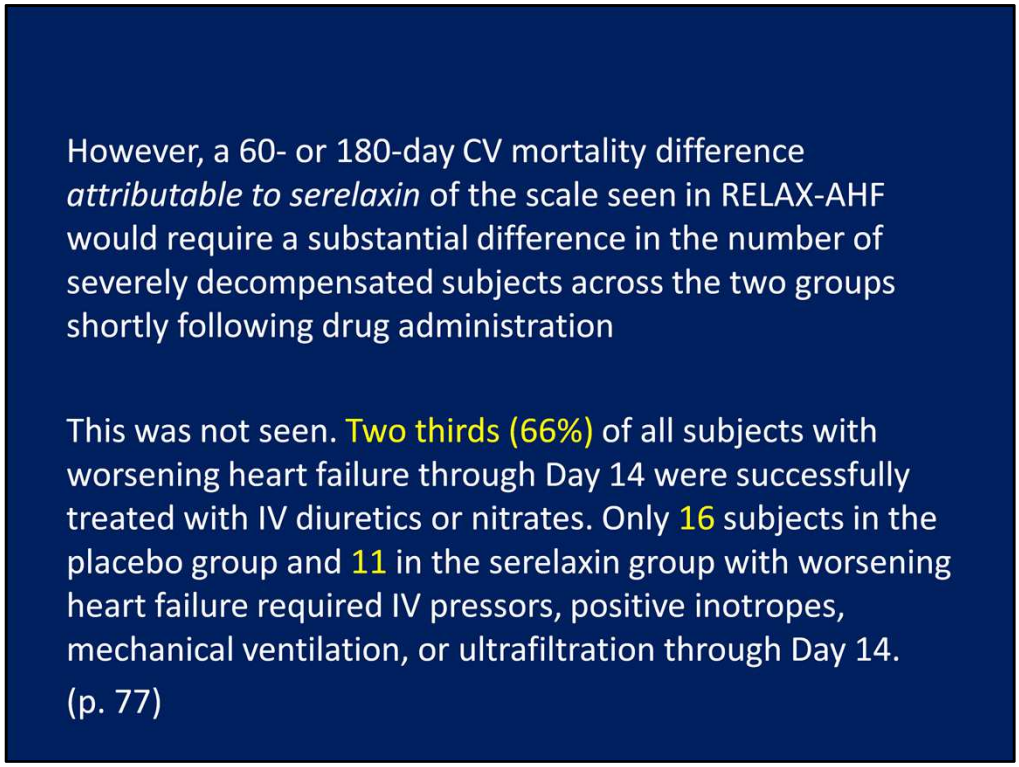
CV mortality finding lacks biological plausibility

- CV mortality at day 180 was neither a primary nor secondary outcome, therefore mortality findings must be investigated in a larger, confirmatory trial before any conclusions can be drawn as to whether this represents a real or chance finding.
- Given that this was a post hoc analysis, the burden is therefore on the sponsor to demonstrate a convincing biological mechanism by which a drug given as a single infusion over 48 hours can result in a mortality benefit weeks and months into the future. For such one-time therapies, one would expect diverging survival curves in the days immediately following therapy, followed by a leveling off of residual effects and **increasingly parallel** curves.

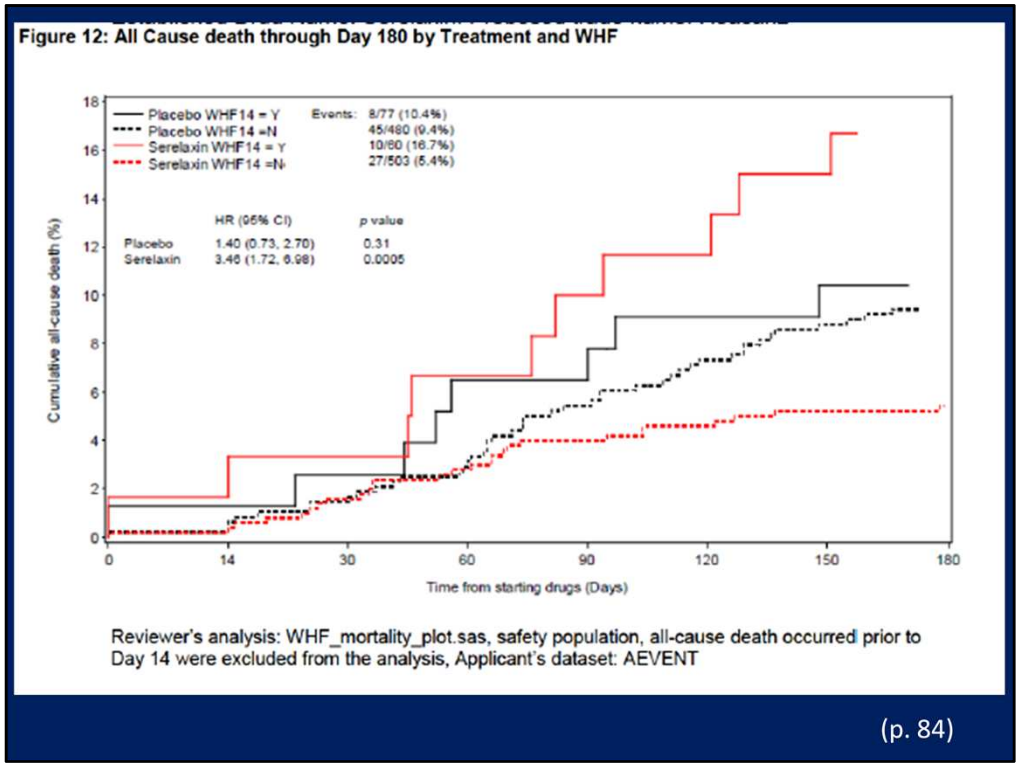
However, a look at the Kaplan-Meier curve, restricted to CV mortality, shows **increasingly divergent** survival curves, something one would expect only with chronic or repeatedly administered acute therapy. (Figure taken from sponsor briefing document, p. 94)



Of course, it could be argued that the initial differences in rates of worsening heart failure would not be reflected in CV mortality differences until a later time point, even as far out as 180 days for patients with severe decompensation and subsequently lengthy ICU stays



FDA Briefing Doc, p. 77. 104 of 157 (66%) subjects with "worsening heart failure" through Day 14 were treated with nothing more than IV diuretics or IV nitrates. Subjects were only counted once, with outcomes considered by decreasing severity (death, followed by ultrafiltration, mechanical ventilation, , IV milrinone, IV pressors, IV nitrates, IV diuretics, and other and finally, rehospitalization for HF).



Furthermore, a cursory look at FDA's post hoc analysis of all-cause mortality rates stratified by "worsening heart failure" status at 14 days makes a causal connection even more unlikely.

As can be seen, virtually all of the 180-day difference in all-cause mortality is driven by the different rates in subjects with *no worsening heart failure* at 14 days. Furthermore, the divergence in rates for these subjects does not begin until approximately *day 60* post-drug administration. Before day 60, the curves were virtually identical. This makes any claim of a causal link between serelaxin administration and all-cause mortality extremely tenuous.

For those subjects with worsening heart failure, placebo subjects actually appeared to have a consistently *lower* rate of all-cause death *throughout* the duration of the 180-day follow-up period. This may be a chance finding (or one that suggests that serelaxin-treated patients who deteriorated were generally sicker to begin with in that they were refractory even to additional [albeit marginally clinically significant] therapy with serelaxin), but is striking nonetheless when considering the overall differences in rates between placebo and serelaxin.

While it is possible that serelaxin was indeed responsible for the CV mortality differences at 60 and 180 days (e.g., through prevention of irreversible cardiac remodeling in the acute phase), until the confirmatory trial currently underway is concluded, the RELAX-AHF mortality findings can only be considered hypothesis-generating and should certainly not factor into today's decision on approval.

Blinding compromised?

- Dr. Blank: *"The serelaxin solution is frothy when shaken and the placebo solution is not. This may have caused unblinding." (p. 40)*
- The fact that both primary endpoints were subjective and that the one purportedly statistically significant outcome was exceedingly marginal (~4 mm on a 100 mm subjective scale) means that even a few unblinded investigators (and patients) may have unintentionally skewed the results in favor of serelaxin

Overly broad indication

- "improve the symptoms of acute heart failure through reduction of the rate of worsening of heart failure"
- The sole symptom measured in either primary outcome was dyspnea *at rest.*
- AHF symptoms are not restricted to dyspnea at rest (particularly for NYHA Class I-III patients):
  - Dyspnea on exertion
  - Peripheral edema
  - Other pulmonary edema/low EF sx (cough, wheezing, fatigue, chest pain, etc…)
- "reduction of the rate of worsening of heart failure" implies an improvement in some objective measure of heart failure severity. No such rigorous objective indicators of heart failure progression (e.g. EF) were systematically measured as part of a prespecified outcome.

Overly broad indication: all AHF patients?

- Wang et al. (2014) compared RELAX-AHF inclusion criteria and study population with 196,770 AHF patients in the Acute Decompensated Heart Failure National Registry-United States (ADHERE-US) and ADHERE-International (ADHERE-I) registries.
- Registries consist of adult inpatients with a primary or secondary discharge diagnosis of acute decompensated heart failure from 2001 to 2009, with consecutive enrollment encouraged.

Wang TS, Hellkamp AS, Patel CB, Ezekowitz JA, Fonarow GC, Hernandez AF. Representativeness of RELAX-AHF Clinical Trial Population in Acute Heart Failure. Circ Cardiovasc Qual Outcomes. 2014 Mar 1;7(2):259-68. doi: 10.1161/CIRCOUTCOMES.113.000418. Epub 2014 Mar 4.

Overly broad indication: all AHF patients?

- Only 16.2% of ADHERE-International and 20.7% of ADHERE-US patients eligible for RELAX-AHF
- RELAX-AHF eligible patients significantly more likely to be older, female, have higher SBP on presentation, more well preserved ejection fraction, better renal function, in addition to a host of other indicators of a markedly healthier AHF population (all $p<0.05$)
- Not surprisingly, therefore: "In-hospital mortality was lower in RELAX-AHF-type than in non–RELAX-AHF-type patients, even after multivariable adjustment (HR 0.59; 95% CI: 0.53–0.66; $P<0.0001$)."

Wang TS, Hellkamp AS, Patel CB, Ezekowitz JA, Fonarow GC, Hernandez AF. Representativeness of RELAX-AHF Clinical Trial Population in Acute Heart Failure. Circ Cardiovasc Qual Outcomes. 2014 Mar 1;7(2):259-68. doi: 10.1161/CIRCOUTCOMES.113.000418. Epub 2014 Mar 4.

"Patients who were considered RELAX-AHF type were ≥18 years of age and met all of the following inclusion criteria: (1) discharge diagnosis of HF, (2) SBP >125 mm Hg, (3) dyspnea at rest or with mild exertion, (4) intravenous diuretic use, and (5) a glomerular filtration rate of 30 to 75 mL/min per 1.73 m2. Patients were excluded if they had intravenous inotropes or vasopressors during their hospital stay or hemoglobin ≤8 g/dL. Patients not meeting all of the above criteria were considered non–RELAX-AHF-type patients."

Overly broad indication: all AHF patients?

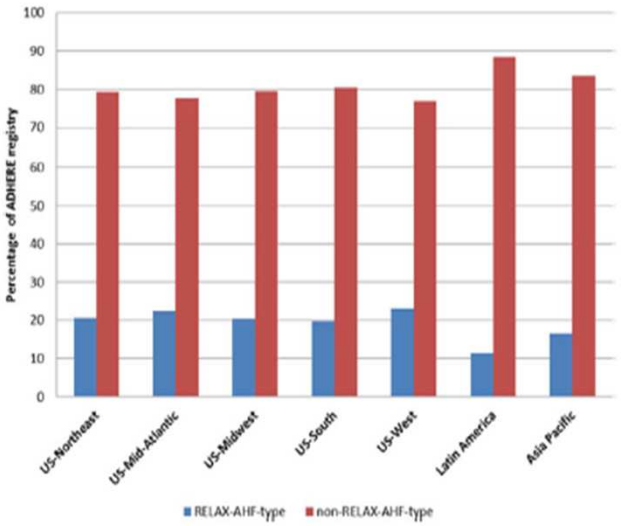


Figure 1. Percentage of Acute Decompensated Heart Failure National Registry (ADHERE) registries that meet or do not meet Relaxin for the Treatment of Acute Heart Failure (RELAX-AHF) entry criteria by region. US indicates United States.

Wang TS, Hellkamp AS, Patel CB, Ezekowitz JA, Fonarow GC, Hernandez AF. Representativeness of RELAX-AHF Clinical Trial Population in Acute Heart Failure. Circ Cardiovasc Qual Outcomes. 2014 Mar 1;7(2):259-68. doi: 10.1161/CIRCOUTCOMES.113.000418. Epub 2014 Mar 4.

Jan 24: EMA rejects serelaxin on identical grounds laid out by FDA's Dr. Blank

"The [EMA] Committee noted that the study results did not demonstrate a benefit for short-term relief of dyspnoea over up to 24 hours, and although some benefit was shown over 5 days it was not clear how this was of clinical relevance.

Furthermore, the Committee had concerns about the way the effectiveness of the medicine in the study had been analysed. The results included calculated values for a number of patients who had died or had required additional treatment for worsening symptoms and whose actual data were not used.

...Since only one main study was included in the application, further studies would be needed to confirm the effectiveness of [serelaxin] in the treatment of acute heart failure.

EMA's explanation of refusal of marketing authorization for serelaxin, 2/21/2014: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/002817/WC500160088.pdf

EMA rejection

"Although the safety of [serelaxin] seemed acceptable, in view of the uncertainties about the benefits of treatment the CHMP was of the opinion, at that point in time, that the benefits of Reasanz did not outweigh its risks and recommended that it be refused marketing authorisation."

Conclusion

- Serelaxin has exceedingly marginal, if any, benefits on subjective, patient-reported dyspnea
- Serelaxin has no proven benefits on any objective cardiovascular outcomes
- Proposed indication bears little resemblance to primary endpoints studied
- Extremely poor generalizability of RELAX-AHF (~20% of all AHF inpatients would have met inclusion criteria)