Author's response to reviews

Title: Outcomes of Patients Hospitalized for Acute Decompensated Heart Failure: Does Nesiritide Make a Difference?

Authors:

Richard J Carroll (richard.carroll@ahss.org)
Zuber D Mulla (zuber.mulla@ttuhsc.edu)
Loran D Hauck (loran.hauck@ahss.org)
Audrey Westbrook (audrey.westbrook@fhmd.org)

Version: 2 Date: 30 August 2007

Author's response to reviews: see over
August 30, 2007

Melissa Norton, MD
Editor-in-Chief

RE: Manuscript 1703742690151101

Dear Dr. Norton:

Thank you for the opportunity to revise our manuscript on nesiritide. Per the reviewers’ suggestions, we have changed the title of our paper to reflect a neutral tone, and we have also made significant changes to the discussion section in order to have a more balanced presentation. We have also deleted any reference to the controversial MEDai severity measure. We now rely on traditional methods and parameters such as the APR-DRG measures.

The number of tables has been reduced from 5 to 4 and the word count has decreased (2753 to 2330). In addition, the paper now complies with the journal’s formatting guidelines. Thank you for your attention.

Sincerely yours,

Zuber D. Mulla, PhD
Associate Professor and Director of Epidemiologic Research
Department of Obstetrics and Gynecology
Texas Tech University Health Sciences Center School of Medicine at El Paso

Adjunct Associate Professor
University of Texas School of Public Health at Houston
El Paso Regional Campus
Original Title: Treatment with Nesiritide is Not Associated with Positive Outcomes in Patients Hospitalized for Acute Decompensated Heart Failure

Reviewer: Jonathan Sackner-Bernstein

The authors’ response is in **bold** font.

---

**Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)**

The authors should provide more information to explain why the outcomes appear worse for mortality when using the MEDai models as opposed to the APR-DRG models. This inconsistency in the mortality effect should be discussed as making this finding less robust, and the two approaches should be compared to the ADHERE registry analytic approaches.

**We have removed any reference to the MEDai severity score. This resulted in the deletion of a table of odds ratios stratified by MEDai, and a revision of the abstract and the references.**

**ADHERE investigators used classification and regression tree analysis (CART) to build a predictive model. As clinicians and epidemiologists, rather than statisticians, we chose to build models based on clinical plausibility rather than p-values.**

MEDai: The authors cite the use of the method in a pneumonia analysis and report on page 8 that the technique was validated. Perhaps this validation methodology and results could be made available as an online supplement. Did it include unique groups of patients for the derivation and validation? Were any of the patient that are part of the analysis set of this paper also part of those derivation and validation sets? Why do the authors want readers to put more weight on the MEDai than the APR-DRG model?

**We have removed any reference to the MEDai severity score throughout the paper. We have now controlled for patient’s race and economic status (Table 2) since these variables are not integrated into the APR-DRG measures but were incorporated into the MEDai severity score.**

The results section portrays the use of nesiritide as one with increased risk of mortality, prolonged length of stay, increased costs and point estimates suggesting even the possibility of increased readmission rates. Yet the authors have not divulged how many events there actually were. By my calculation, there were only 13 deaths in the nesiritide treated population. Such a small number means that errors are possible which may be reflected in the results being somewhat different using the two methodologies for mortality risk estimates. Can the authors provide the number of events?

**The reviewer has calculated that we had 13 deaths in our sample. The actual number is larger. In the first sentence of our Results section in our original draft we did provide the overall number of deaths: 805. We have now added the following sentence to the Results section after the sentence that states the 805 deaths: The distribution of the deaths was as follows: 126 patients who received nesiritide, and 679 patients who did**
not receive nesiritide (Table 1). In addition, we have added the numbers for all of the variables listed in Table 1. Previously we had just listed percents. We have also added the number of events/outcomes (hospital death, prolonged length of stay, elevated pharmacy cost, and readmission within 31 days for cardiac condition) to Table 1. A few sentences (including the one discussing Figure 1) have been rearranged in the first paragraph of the Results section to make this section flow better.

In the second paragraph of our revised Results section we have also added more detail (new text is underlined) on the crude measures of effect (relative risk and odds ratio) that quantify the association between nesiritide and hospital death.

The authors state an interpretation of the study findings in the title. There are reasons to advocate for a different tact. First, the tone is one where it seems that the paper will “prove” this to be the case, when some could argue that this analytic approach cannot be claimed as proof. Second, the analyses are also relevant for showing that nesiritide is not so clearly associated with risk. This manuscript provides justification for continuing enrollment in the ongoing clinical outcomes trial assessing nesiritide, as one can point to these data as supporting that the trial is ethical.

We have changed the title (which is strong and presents a conclusion) to a neutral title: Outcomes of Patients Hospitalized for Acute Decompensated Heart Failure: Does Nesiritide Make a Difference?

-------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

On page 11, the authors focus on the decreased use of nesiritide in clinical practice. Citing the paper by Hauptman et al from JAMA would seem ideal.

When the authors refer to the meta-analysis from Sackner-Bernstein et al, it would be best to also cite the follow-up analysis also published in JAMA that includes more data and provides a more accurate estimate of risk.

We have now cited this follow-up analysis (the JAMA research letter by Aaronson and Sackner-Bernstein). It is new reference number 22. Results from this study are discussed on the top of page 11 of our paper (underlined text).

-------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

On page 14, the authors state that nesiritide improves dyspnea. In fact, review of the data from VMAC will show that this is somewhat of a marginal effect from both a statistical and especially clinical perspective. One could argue whether this should be stated with the same conviction as the hemodynamic effect, merely because the approved labeling says them in this fashion.
Reviewer: John Rumsfeld

General

The topic of this study is of high interest, and the study is generally well-conducted. Specific limitations include the single-center setting, and primarily that therapeutic evaluations using observational data can yield associations but cannot determine causality. The authors over-state the findings, and the manuscript should be re-written to be much more circumspect in this regard, as there is high potential for both selection bias and unmeasured confounding.

The reviewer states above that this is a “…single-center…” study. Actually the study involved 31 hospitals in multiple states. This was stated in the original draft in the first sentence of the section entitled Patient Population.

We have toned down our interpretation by deleting verbiage including the following text in the Discussion:

“Our unadjusted statistics are certainly disturbing: (1) a 59% increased risk of mortality, (2) a 141% increase in the risk of prolonged length of stay, and (3) a 524% increase in the odds of having an increased pharmacy cost. These results provide an incomplete picture without severity adjustment, particularly given the differences in the frequency distributions of severity levels in our data.”

Our adjusted hospital mortality odds ratios were not statistically significant and so we also deleted text in the Discussion that referred to our mortality results as “discomforting.”

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. The principal limitations of this study (as with all observational evaluations of therapeutic efficacy) are selection bias/confounding by indication (the nesiritide patients were much sicker, shown clearly by the patient characteristics/severity of disease measurements provided), and potential for unmeasured confounding (almost certainly the patients who did and didn’t receive nesiritide differed by measures not available in this study). This does not invalidate the observational associations found, and there can still be some value in reporting such observational data as a reflection of ‘real-world’ practice. However, strongly recommend that the authors go through the paper and a) remove all language that implies causality (e.g. ‘significantly increased the odds’, ‘nesiritide remained a significant risk factor’, ‘increase in the odds’, ‘were linked to nesiritide treatment’; b) substitute language that reports associations only (e.g. ‘was associated with higher odds’, etc.); c) consider a propensity matched analysis utilizing all available baseline data on patient demographics, clinical history and comorbidities, and severity of disease – note that this will not obviate the potential for selection bias/unmeasured confounding, but would be more robust than the current risk adjustment; and d) be much more circumspect with regard to implications/discussion….in particular, the Discussion should be toned down –
sicker patients will have longer length of stay, cost more, and have worse outcomes….since these patients were also more likely to receive nesiritide the deck was stacked; this must be acknowledged, and the Discussion should note that nesiritide is being used in sicker patients, who have longer LOS, cost more, and have worse outcomes – then, as already is in the final paragraph – make the call for additional randomized trials evaluating the efficacy and cost effectiveness of nesiritide.

The reviewer suggests that we, “…consider a propensity matched analysis…” Propensity scores are especially indicated when there are many covariates in the model but there are too few events (please see page 268 of the text Epidemiologic Methods: Studying the Occurrence of Illness by Koepsell and Weiss published in 2003). We evaluated four outcomes. The outcome with the fewest number of events was hospital mortality. We observed 805 deaths in the hospital. Dividing the number of deaths by 10 yields a figure of 80.5. This means a logistic regression model where hospital death is the outcome can accommodate 80 independent variables (e.g., nesiritide plus 79 medication variables). Please refer to pp. 346-347 of the text Applied Logistic Regression 2nd edition by Hosmer and Lemeshow for a description of this sample size rule which calls for a minimum of 10 events per parameter. At most our models included only 15 independent variables: nesiritide, 9 medication variables, an APR-DRG variable, three dummy (indicator) variables for race/ethnicity, and low economic status. Therefore we feel that an analysis with propensity scores is not indicated.

For the reader, propensity scores add another layer of complexity to the analysis since you must model logit (nesiritide) = covariates and then calculate the expected probabilities of being exposed to nesiritide and then return to the original research question of logit(hospital death) = nesiritide and either match on the propensity score or add it to the logistic regression model. A more direct approach is to follow the traditional method of adjustment, which we did: logit(hospital death) = nesiritide + confounders.

2. Methods: How much missing data was there….and thus, how many patients were excluded from analyses on this basis? If a substantial number and missing data rates relatively low, why was multiple imputation not utilized?

Only 613 patient records were excluded. To clarify, only 2.4% of the original dataset which contained 25,943 records had to be excluded because of missing values for the variables under study.

Imputation is not required in this situation due to the small numbers of missing values and the wariness that some readers have towards imputation. To clarify, many of the current multiple imputation methods assume data are missing at random (MAR) [Potthoff RF et al. Stat Methods Med Res, June 2006]; however, there is no direct method to test the assumption of MAR [Potthoff RF et al. Stat Methods Med Res, June 2006]. Potthoff and colleagues [Stat Methods Med Res, June 2006] state, “There are limits to the ability of sophisticated statistical methods to correct for missing data.” Again, we do not want to alienate any readers. Prof. Frank E. Harrell, Jr., in his text Regression Modeling Strategies (page 44) discusses imputation and prefaxes this topic by writing, “Many nonstatisticians find the notion of estimating data distasteful…”

3. Strongly recommend a secondary analysis (or switch this to the primary analysis) utilizing unique patients (18298 instead of 25,330) as the use of repeat
admissions likely reflects a clustering of sicker patients (wish associated bias toward more nesiritide use).

As stated in our methods section, we accounted for repeated admissions using generalized estimating equations. This topic is discussed in the original version under the subheading of Repeated Measurements.

4. Why was PROC GENMOD used instead of PROC LOGISTIC for the multivariable analyses, which appear to have been multiple logistic regression?

Yes, we did use multiple logistic regression to analyze the data which was stated at the beginning of paragraph 3 of the original Statistical Analysis section. This paragraph states that we used PROC GENMOD. The reason why PROC GENMOD was used to perform logistic regression rather than PROC LOGISTIC is because PROC LOGISTIC does not allow the user to use generalized estimating equations (GEE). GEE is a standard technique to account for correlated outcome data that may arise because of repeated measurements or the clustering or nesting of subjects (please see the text Logistic Regression using the SAS System: Theory and Application by Paul D. Allison).

5. Given the significant baseline differences between the nesiritide and no-nesiritide patients, there is no reason to present (and then reiterate in the Discussion) the unadjusted results. Somewhat related, any additional clinical characteristics (demographic, clinical history and comorbidities, and hospital course in addition to what is provided in the manuscript) available to the authors should be added to Table 1 to better show the ways in which the 2 cohorts differ (and these variables should be utilized in adjustment as noted in earlier comment), with statistical comparison.

We respectfully submit that the unadjusted results are valuable since they show the degree of confounding. If the crude (unadjusted) and adjusted odds ratios differ by 10 or more percent then confounding is traditionally considered to be present (please refer to Sander Greenland’s paper in the American Journal of Public Health 1989;79(3):340-9). Only 4 unadjusted results are presented (found in the first row of Table 2) so they do not take up much space.

We agree with the reviewer that additional baseline data should be shown in Table 1 and additional variables should be controlled for so we have added the median patient ages in a footnote to Table 1 and added the prevalence of low economic status and race/ethnicity to Table 1. Furthermore, several of the odds ratios in Table 2 are now also adjusted for race and low economic status. Age is incorporated into the two APR-DRG severity measures by the 3M company and therefore the age term was not included in our regression models.