Evaluating Treatment Efficacy by Multiple End Points in Phase II Acute Heart Failure Clinical Trials

Analyzing Data Using a Global Method

Hengrui Sun, MD, MPH; Beth A. Davison, PhD; Gad Cotter, MD; Michael J. Pencina, PhD; Gary G. Koch, PhD

**Background**—To assess concomitant simultaneous effects on multiple end points using global statistical methods in phase II acute heart failure studies.

**Methods and Results**—Using simulations we have assessed different statistical methods to evaluate concomitant effects of a new intervention on dyspnea relief (using 2 measures), length of hospital stay, worsening heart failure to 5 days, mortality, and heart failure readmission to 30 days. Treatment effect scenarios included large (20% to 28% relative improvements) and very large (30% to 43% relative improvements) effects among others. Placebo responses and correlations among end points typical in recent acute heart failure clinical trials were used. Powers for the average Z score exceeded 70% with ≥75 patients per group for 35% relative improvement across all 6 end points. Assessing dyspnea alone generally provides lower power than the average Z score approach, with power deducted ≥50% under most of scenarios. Other approaches generally provide lower power than the average Z score method.

**Conclusions**—Assessing the effects of new therapies on multiple clinical end points using the average Z score enables detection of therapeutic efficacy using sample sizes of 100 to 150 patients per group, approximately double the power achievable assessing the effects on dyspnea alone. (Circ Heart Fail. 2012;5:742–749.)

**Key Words:** heart failure ■ clinical trial ■ statistics ■ end points ■ global method ■ power

**Clinical Perspective on p 749**

Recently, several methods that focus on combining multiple end points into a single end point have been discussed in scientific journals. Felker et al proposed a global rank end point. Finkelstein-Schoenfeld’s test is based on a hierarchy of the end points. A clinical composite outcome of worsened/not changed/improved has frequently been used in cardiovascular trial design. With correlated end points, the average Z score, which...
is an extension of O'Brien's rank-sum procedure, is of interest because it is simple to implement and potentially powerful. Another choice, which does not require the combining of the end points, is to evaluate each end point separately, and then define a trial as success if a certain criterion is met. A single clinical end point (dyspnea) is also compared which in previous studies was found to be the most sensitive to treatment.

In this study, we assess the ability of the above-mentioned approaches in detecting a treatment effect by simultaneous assessment of multiple clinically relevant end points pertaining to symptom relief and outcomes, and propose the best method in terms of power for smaller sample sizes.

### Methods

#### End Points Selection

The main treatment targets for AHF are improving symptoms and reducing mortality; reducing the cost of treatment is also a consideration. The following end points reflect these treatment goals: all-cause death through day 30 (mortality); rehospitalization because of heart failure to day 30 and time to worsening heart failure to day 5 (both pertain to disease recurrence and recurrence of symptoms, rates of re-admission also pertain to economical cost to society); dyspnea visual analog scale (VAS) score area under the change from baseline curve (AUC) from baseline to day 5 and moderately or markedly better dyspnea at 6, 12, and 24 hours by Likert scale (both pertain to rapidity of symptom improvement during the initial treatment period); and length of initial hospital stay (pertaining to disease burden on the individual and economical cost to society). In recent AHF trials, dyspnea was measured using 2 scales: absolute intensity on a 10-cm VAS,8,10 and improvement relative to baseline on a Likert scale from −3 (markedly worse) to +3 (markedly better).6,8,11

#### Size of the Treatment Effect and Scenarios

On the basis of our clinical experience and observations from previous AHF trials, sets of treatment effects were assumed for each of the 6 end points (Table 1). The proposed treatment effects assume that improving symptoms or reducing outcomes by 15% to 20% represents moderate acceptable beneficial effects. Effects smaller than that are likely to be less significant clinically, and larger effects are highly clinically significant. Standardized effect sizes are also computed and presented with the clinical effects, so that the magnitude of each clinical effect can be compared across end points. By assigning different combinations of effect sizes for the 6 end points, hypothetical scenarios were designed to mimic situations that one could expect to observe in an AHF phase II study (Table 2). Null treatment effects (scenario 15) are included to evaluate the type I error (false positive) rate for each method.

### Simulation

Six end points were simulated to represent death to day 30, heart failure rehospitalization to day 30, time to worsening heart failure to day 5, dyspnea VAS AUC from baseline to day 5, moderately/markedly better dyspnea at 6, 12, and 24 hours by Likert, and length of initial hospital stay. On the basis of previous data, we assumed moderate correlations among the end points (Table 3). Multivariate normal random variables were generated using an initial working correlation structure, which was then iteratively modified to achieve the desired correlations for the 6 end points. Time-to-event end points were derived by transformations to assumed exponential distributions with constant hazard rates; dichotomous end points were derived based on categorizations so that binomial distributions applied. A total of 1000 trials were simulated for each treatment effect scenario with sample sizes of 50, 75, 100, 125, and 150 patients per group for each trial. Trials for independent end points were also simulated in parallel to compare with the powers from the correlated data.

### Statistical Analysis

Under different treatment effect scenarios and sample sizes, the powers of 5 methods of combining these 6 end points were compared.

#### Average Z Score

O'Brien proposed that simple or weighted summation criteria for single end points with standard procedures lead to asymptomatically normal statistics. Here, we extended this approach to include continuous, dichotomous, and time-to-event end points. Specifically, continuous and dichotomous variables were converted to z scores by subtracting the mean of the pooled groups from an individual's value and dividing by the standard deviation of the pooled group; time-to-event variables were first transformed to log-rank scores and then converted to standard Z scores by subtracting the mean and dividing by the standard deviation of the pooled data. Z scores for all end points were then aligned so that bigger scores represent better outcomes. Averaging Z scores across end points for each subject, the treatment groups were compared using a Wilcoxon rank-sum test.

#### Global Rank

Feldker et al proposed a hierarchical global rank outcome for use in AHF studies encompassing mortality, symptoms, and biomarkers. We adapted the approach to exclude biomarkers and include clinical outcomes representing treatment targets in AHF. Subjects were ranked hierarchically in broad categories 1 to 6, and then more finely within these broad categories, with worse outcomes having lower ranks. Details are shown in the Figure. Treatment groups were compared with respect to the global rank outcome with a Wilcoxon rank-sum test.

### Finkelstein-Schoenfeld Hierarchical Composite

Finkelstein and Schoenfeld's method is a modification of the generalized Wilcoxon test, and was described for combining mortality and longitudinal measures. We have adapted this approach to comprise the 6 AHF end points. Generally, each patient in the active treatment

---

**Table 1. Magnitude of the Clinical Effects With Corresponding Standardized Effect Sizes for Each End Point**

<table>
<thead>
<tr>
<th>Response in Active Group (Standardized Effect Size)</th>
<th>Placebo</th>
<th>Small Negative Clinical Effect</th>
<th>Small Clinical Effect</th>
<th>Medium Clinical Effect</th>
<th>Large Clinical Effect</th>
<th>Very Large Clinical Effect</th>
<th>Large+Clinical Effect</th>
<th>Modified Large+Clinical Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death 30 d (event rate)</td>
<td>0.07</td>
<td>0.077 (−0.028)</td>
<td>0.063 (0.028)</td>
<td>0.06 (0.041)</td>
<td>0.055 (0.062)</td>
<td>0.047 (0.098)</td>
<td>0.051 (0.080)</td>
<td>0.05 (0.084)</td>
</tr>
<tr>
<td>Rehospital HF 30 d (event rate)</td>
<td>0.14</td>
<td>0.156 (−0.045)</td>
<td>0.125 (0.044)</td>
<td>0.115 (0.075)</td>
<td>0.11 (0.091)</td>
<td>0.097 (0.13)</td>
<td>0.104 (0.11)</td>
<td>0.115 (0.075)</td>
</tr>
<tr>
<td>WHF 5 d (event rate)</td>
<td>0.2</td>
<td>—</td>
<td>0.18 (0.051)</td>
<td>0.17 (0.077)</td>
<td>0.155 (0.12)</td>
<td>0.135 (0.17)</td>
<td>0.145 (0.15)</td>
<td>0.16 (0.10)</td>
</tr>
<tr>
<td>VAS AUC 5 d (mean)</td>
<td>1680</td>
<td>—</td>
<td>1850 (0.063)</td>
<td>2000 (0.12)</td>
<td>2150 (0.17)</td>
<td>2400 (0.27)</td>
<td>2300 (0.23)</td>
<td>2500 (0.30)</td>
</tr>
<tr>
<td>Likert 6, 12, 24 h (proportion)</td>
<td>0.23</td>
<td>—</td>
<td>0.255 (0.058)</td>
<td>0.27 (0.092)</td>
<td>0.29 (0.14)</td>
<td>0.325 (0.21)</td>
<td>0.31 (0.18)</td>
<td>0.34 (0.25)</td>
</tr>
<tr>
<td>Length of initial hospital stay (mean day)</td>
<td>10</td>
<td>—</td>
<td>9.0 (0.14)</td>
<td>8.5 (0.22)</td>
<td>8.0 (0.29)</td>
<td>7.0 (0.43)</td>
<td>7.5 (0.36)</td>
<td>8.5 (0.22)</td>
</tr>
</tbody>
</table>

HF indicates heart failure; WHF, worsening heart failure; VAS, visual analog scale; AUC, area under the change from baseline curve.
group is compared with each patient in the placebo group in a pair-wise manner. The hierarchy of the end point was formed as for the global rank method. Within each category, a score 1, 0, or −1 is assigned to the subject if the active treatment group did better, same or worse than the patient in the control group. The test statistic is based on the sum of net scores for all treated subjects.

### Clinical Composite Outcome of Worsened/Not Changed/Improved

Each patient is assigned an outcome of worsened, not changed, or improved based on criteria using the 6 end points, with scores of 0, 1, and 2 representing the 3 outcomes, respectively. Success was defined as patient-reported moderate or markedly better dyspnea, or VAS AUC ≥ 936 mm-h (average of 8 mm improvement) in the

| Table 2. Treatment Effect Scenarios Representing Plausible Clinical Trial Results |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Scenario | Description | Death 30 d | Rehospital HF 30 d | WHF5d | VAS AUC 5 d | Likert 6, 12, 24 h | Length of Initial Hospital Stay |
| 1 | Modest effects on 2 dyspnea end points and LOS, “trends” on other end points | Small | Small | Small | Medium | Medium | Medium |
| 2 | Modest effects on all 6 end points | Medium | Medium | Medium | Medium | Medium | Medium |
| 3 | Large effects on 2 dyspnea end points, modest effect on LOS, “trends” on other end points | Small | Small | Small | Large | Large | Large |
| 4 | Large effects on 2 dyspnea end points and LOS, modest effect on other end points | Medium | Medium | Medium | Large | Large | Large |
| 5 | Large effects on all 6 end points | Large | Large | Large | Large | Large | Large |
| 6 | Large effects on 2 dyspnea end points, modest effect on LOS, small effect on WHF 5 d, neutral on death or rehospitalization | No effect | No effect | Small | Large | Large | Large |
| 7 | Large effects on 2 dyspnea end points, modest effect on LOS and WHF 5 d, small negative effect on death or rehospitalization | Neg small | Neg small | Medium | Large | Large | Medium |
| 8 | Very large effects on 2 dyspnea end points, modest effect on LOS and WHF 5 d, small negative effect on death or rehospitalization | Neg small | Neg small | Medium | Very large | Very large | Medium |
| 9 | Trends in 2 dyspnea end points and LOS, modest effect on other end points | Medium | Medium | Medium | Small | Small | Small |
| 10 | “Trends” in 2 dyspnea end points and LOS, modest effect on WHF 5 d, large effect on death or rehospitalization | Large | Large | Medium | Small | Small | Small |
| 11 | Very large effect on death and rehospitalization, neutral on other end points | Very large | Very large | No effect | No effect | No effect | No effect |
| 12 | Large + effects on all 6 end points | Large + | Large + | Large + | Large + | Large + | Large + |
| 13 | Modified large + effects on all 6 end points | Modified | Modified | Modified | Modified | Modified | Modified |
| 14 | Very large effects on all 6 end points | Very large | Very large | Very large | Very large | Very large | Very large |
| 15 | Neutral on all 6 end points | No effect | No effect | No effect | No effect | No effect | No effect |

| Table 3. Targeted Correlation Structure Among 6 End Points (Pearson Correlations. Death 30 d and Rehospital HF 30 d Were Transformed to Log-Rank Scores Before Calculating the Correlations) |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Description | Death 30 d | Rehospital HF 30 d | WHF5 d | VAS AUC 5 d | Likert 6, 12, 24 h | Length of Initial Hospital Stay |
| Death 30 d | 1 | 0.05 | 0.15 | 0.05 | 0 | 0 |
| Rehospital HF 30 d | 0.05 | 1 | 0.05 | 0 | 0 | 0 |
| WHF5 d | 0.15 | 0.05 | 1 | 0.65 | 0.15 | 0.2 |
| VAS AUC 5 d | 0.05 | 0 | 0.65 | 1 | 0.25 | 0.2 |
| Likert 6, 12, 24 h | 0 | 0 | 0.15 | 0.25 | 1 | 0.05 |
| Length of initial hospital stay | 0 | 0 | 0.2 | 0.2 | 0.05 | 1 |

HF indicates heart failure; VAS, visual analog scale; AUC, area under the change from baseline curve; LOS, length of initial hospital stay.
absence of any criterion for failure. Failure was defined as the occurrence of any of the following: death or heart failure readmission through day 30, or worsening heart failure by day 5. Patients were classified as having unchanged treatment status if they met neither the treatment success nor treatment failure criteria. A Cochran-Mantel-Heanszel test with row mean score is used to determine whether treatment is effective.

**P Value Criteria**

Each end point was assessed individually, with treatment groups compared with respect to time-to-event end points by log-rank test, dichotomous end points by \( \chi^2 \) test, and continuous end points by Wilcoxon rank-sum test. Two criteria to define treatment success were evaluated: (1) the treatment effect was in a favorable direction with a 1-sided \( P \) value \( \leq 0.2 \) for \( \geq 4 \) of the end points, and none of the end points had a treatment effect in the opposite direction with a 1-sided \( P \) value \( \leq 0.2 \); and (2) 2 of the 6 end points have 1-sided \( P \) values \( \leq 0.1 \) with 4 of the 6 1-sided \( P \) values \( \leq 0.2 \) and all 1-sided \( P \) values \( \leq 0.8 \).

All the analyses were performed using SAS release 9.2 (SAS Institute, Inc., Cary, NC).

**Results**

Table 4 illustrates the powers for each approach under variant scenarios for correlated end points. The power was calculated as the proportion of successful trials among 1000 simulated trials. For the \( P \) value criteria, a successful trial was defined as above; for the other approaches, an individual trial was declared a success if the 1-sided \( P \) value \( \leq 0.025 \), in the direction favoring active treatment. With end points correlated, powers are generally lower than with independent end points (supplementary data) for the average \( Z \) score and clinical composite outcome approaches; for the global rank and Finkelstein-Schoenfeld methods, correlation of the end points does not affect power much, with some higher and some lower than the power for independent data; the \( P \) value criteria generally have larger power when end points are correlated.

The average \( Z \) score provides the highest power compared with the other approaches in all treatment scenarios except when large effects on mortality and morbidity and neutral effects on other end points were assumed (scenario 11), where the hierarchical methods (global rank or Finkelstein-Schoenfeld) are slightly better. Powers for the average \( Z \) score exceed 70% with large effects on all 6 end points, with \( \geq 150 \) patients per treatment group for scenario 5 (20% to 28% relative improvement), 100 subjects per group for scenario 12 (25% to 37% relative improvement), 125 per group for scenario 13 (15% to 49% relative improvement), and with 75 per group for scenario 14 (30% to 43% relative improvement). The global rank outcome provides 70% power for scenario 14 with a sample size of \( \geq 150 \) per group; the Finkelstein-Schoenfeld method with a sample size of \( \geq 150 \) per group for scenario 12, and with \( 100 \) per group for scenario 14; and both \( P \) value criteria with a sample size of \( \geq 150 \) per group for scenario 12, and 125 per group for scenario 14.

The single end point (\( \text{VAS AUC 5 day} \)) generally provides lower power compared with the average \( Z \) score method except in scenario 8 with small negative effects in morbidity and mortality and very large effect for dyspnea. The highest power it reaches is about 70% in scenario 13 with sample size 150 per group; and power at 60% in scenarios 8, 13, and 14 with sample size 150, 125, and 150 per group respectively, which represents the large + and very large effect. With zero or slightly negative effects on mortality and morbidity end points, all 5 methods for combining end points lose substantial power.

All the approaches have reasonable false positive rates (scenario 15) for correlated data, 1-sided around 2.5%. With independent end points, both \( P \) value criteria have low false positive rates (\( \approx 1\% \)), whereas the other approaches have similar results for this as for correlated data.

**Discussion**

Because clinically meaningful end points such as dyspnea improvement and readmission or death occur rarely or have a low signal-to-noise ratios of effect size to standard deviation, larger sample size are usually required to see significant effects. Although phase II studies are relatively small in size by design, expecting significance at the traditional 2-sided 0.05 level for each individual clinical end point can be unrealistic. One approach would be to analyze each end point separately, presenting multiple \( P \) values and an overall conclusion. However, the increasing of overall type I error and where to draw the line between go and no go need to be considered carefully. The other approach, which may be more practical and allow structured decisions based on predetermined cut offs, is to combine all the end points to form a single end point. Thus, by assessing this single end point, we evaluate several aspects of the treatment response.
Table 4. Power Comparison at Variant Scenario and Sample Size Combinations by Different Methods. Data Were Simulated With Prespecified Correlation Structure

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description</th>
<th>Sample Size (n per Group)</th>
<th>Average Z Score</th>
<th>Global Rank</th>
<th>Finkelstein-Schoenfeld Hierarchical Composite</th>
<th>Clinical Composite</th>
<th>P value Criteria</th>
<th>Single End Point–VAS AUC 5 d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Modest effects on 2 dyspnea end points and LOS, “trends” on other end points</td>
<td>50</td>
<td>0.142</td>
<td>0.067</td>
<td>0.091</td>
<td>0.047</td>
<td>0.113</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.206</td>
<td>0.104</td>
<td>0.128</td>
<td>0.065</td>
<td>0.141</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.254</td>
<td>0.11</td>
<td>0.156</td>
<td>0.078</td>
<td>0.186</td>
<td>0.171</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.302</td>
<td>0.138</td>
<td>0.201</td>
<td>0.104</td>
<td>0.226</td>
<td>0.206</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.368</td>
<td>0.152</td>
<td>0.211</td>
<td>0.103</td>
<td>0.255</td>
<td>0.24</td>
</tr>
<tr>
<td>2</td>
<td>Modest effects on all 6 end points</td>
<td>50</td>
<td>0.161</td>
<td>0.094</td>
<td>0.116</td>
<td>0.071</td>
<td>0.138</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.24</td>
<td>0.136</td>
<td>0.164</td>
<td>0.098</td>
<td>0.185</td>
<td>0.171</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.302</td>
<td>0.157</td>
<td>0.205</td>
<td>0.119</td>
<td>0.235</td>
<td>0.213</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.36</td>
<td>0.192</td>
<td>0.254</td>
<td>0.153</td>
<td>0.296</td>
<td>0.273</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.424</td>
<td>0.205</td>
<td>0.285</td>
<td>0.17</td>
<td>0.33</td>
<td>0.315</td>
</tr>
<tr>
<td>3</td>
<td>Large effects on 2 dyspnea end points, modest effect on LOS, “trends” on other end points</td>
<td>50</td>
<td>0.176</td>
<td>0.09</td>
<td>0.116</td>
<td>0.054</td>
<td>0.15</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.272</td>
<td>0.123</td>
<td>0.165</td>
<td>0.071</td>
<td>0.191</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.322</td>
<td>0.153</td>
<td>0.198</td>
<td>0.091</td>
<td>0.233</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.398</td>
<td>0.183</td>
<td>0.258</td>
<td>0.124</td>
<td>0.281</td>
<td>0.261</td>
</tr>
<tr>
<td>4</td>
<td>Large effects on 2 dyspnea end points and LOS, modest effect on other end points</td>
<td>50</td>
<td>0.256</td>
<td>0.118</td>
<td>0.16</td>
<td>0.082</td>
<td>0.206</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.363</td>
<td>0.161</td>
<td>0.232</td>
<td>0.11</td>
<td>0.273</td>
<td>0.261</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.439</td>
<td>0.207</td>
<td>0.295</td>
<td>0.146</td>
<td>0.324</td>
<td>0.309</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.553</td>
<td>0.258</td>
<td>0.348</td>
<td>0.181</td>
<td>0.387</td>
<td>0.376</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.649</td>
<td>0.288</td>
<td>0.42</td>
<td>0.201</td>
<td>0.428</td>
<td>0.417</td>
</tr>
<tr>
<td>5</td>
<td>Large effects on all 6 end points</td>
<td>50</td>
<td>0.291</td>
<td>0.147</td>
<td>0.194</td>
<td>0.114</td>
<td>0.246</td>
<td>0.228</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.421</td>
<td>0.205</td>
<td>0.293</td>
<td>0.155</td>
<td>0.342</td>
<td>0.323</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.504</td>
<td>0.263</td>
<td>0.347</td>
<td>0.203</td>
<td>0.404</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.618</td>
<td>0.326</td>
<td>0.415</td>
<td>0.257</td>
<td>0.479</td>
<td>0.461</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.712</td>
<td>0.371</td>
<td>0.508</td>
<td>0.303</td>
<td>0.545</td>
<td>0.532</td>
</tr>
<tr>
<td>6</td>
<td>Large effects on 2 dyspnea end points, modest effect on LOS, small effect on death or rehospitalization</td>
<td>50</td>
<td>0.152</td>
<td>0.056</td>
<td>0.075</td>
<td>0.038</td>
<td>0.119</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.227</td>
<td>0.092</td>
<td>0.105</td>
<td>0.054</td>
<td>0.144</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.275</td>
<td>0.106</td>
<td>0.136</td>
<td>0.066</td>
<td>0.186</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.337</td>
<td>0.11</td>
<td>0.164</td>
<td>0.071</td>
<td>0.211</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.406</td>
<td>0.126</td>
<td>0.184</td>
<td>0.073</td>
<td>0.231</td>
<td>0.219</td>
</tr>
<tr>
<td>7</td>
<td>Large effects on 2 dyspnea end points, modest effect on LOS and WHF 5 d, small negative effect on death or rehospitalization</td>
<td>50</td>
<td>0.138</td>
<td>0.043</td>
<td>0.049</td>
<td>0.032</td>
<td>0.096</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.204</td>
<td>0.075</td>
<td>0.079</td>
<td>0.049</td>
<td>0.116</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.245</td>
<td>0.071</td>
<td>0.097</td>
<td>0.057</td>
<td>0.154</td>
<td>0.133</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.293</td>
<td>0.08</td>
<td>0.116</td>
<td>0.056</td>
<td>0.178</td>
<td>0.171</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.348</td>
<td>0.089</td>
<td>0.128</td>
<td>0.058</td>
<td>0.182</td>
<td>0.173</td>
</tr>
<tr>
<td>8</td>
<td>Very large effects on 2 dyspnea end points, modest effect on LOS and WHF 5 d, small negative effect on death or rehospitalization</td>
<td>50</td>
<td>0.215</td>
<td>0.066</td>
<td>0.087</td>
<td>0.039</td>
<td>0.13</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.313</td>
<td>0.105</td>
<td>0.132</td>
<td>0.065</td>
<td>0.167</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.376</td>
<td>0.118</td>
<td>0.149</td>
<td>0.065</td>
<td>0.2</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.481</td>
<td>0.136</td>
<td>0.195</td>
<td>0.077</td>
<td>0.236</td>
<td>0.231</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.561</td>
<td>0.151</td>
<td>0.212</td>
<td>0.077</td>
<td>0.23</td>
<td>0.227</td>
</tr>
<tr>
<td>9</td>
<td>“Trends” in 2 dyspnea end points and LOS, modest effect on other end points</td>
<td>50</td>
<td>0.091</td>
<td>0.048</td>
<td>0.058</td>
<td>0.043</td>
<td>0.081</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.111</td>
<td>0.076</td>
<td>0.079</td>
<td>0.051</td>
<td>0.103</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.143</td>
<td>0.079</td>
<td>0.104</td>
<td>0.069</td>
<td>0.128</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.165</td>
<td>0.097</td>
<td>0.122</td>
<td>0.086</td>
<td>0.148</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.183</td>
<td>0.101</td>
<td>0.134</td>
<td>0.087</td>
<td>0.164</td>
<td>0.14</td>
</tr>
</tbody>
</table>

(Continued)
simultaneously, and a P value for the treatment difference with respect to the combined end point represents a formal test of the hypothesis.

Because the average Z score takes the average across the end points, it is apparently more powerful than the single end point (dyspnea VAS AUC at day 5), even when slightly negative results for the 2 morbidity and mortality end points and a large effect on dyspnea are expected. For this scenario, the overall positive result outweighs the negative result, thus higher powers are expected. The global rank and Finkelstein-Schoenfeld methods place priority in the higher order part of the hierarchy. Without apparent beneficial effect on the first 2 outcomes, the effect of outcomes lower in the hierarchy will not show much importance. Therefore, they perform worse than the single end point with zero or slightly negative effects on the 2 morbidity and mortality end points. The 2 P value criteria generally perform better than the single end point except in the 2 scenarios where slightly negative effects exist. Under all scenarios, there are no signs that the clinical composite is any better than the single dyspnea end point.

The global rank and Finkelstein-Schoenfeld hierarchical composite are similarly constructed. They both place hierarchical order to the outcomes, emphasizing higher order end points, which are death and rehospitalization in this case, while putting feeling better to a relatively less important position. Because in small phase II studies the rates of morbidity and mortality are low and the confidence intervals around the point estimates are large, an apparent excess of events in the active treatment group that is likely a chance

Table 4. Continued

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description</th>
<th>Sample Size (n per Group)</th>
<th>Average Z Score</th>
<th>Global Rank</th>
<th>Finkelstein-Schoenfeld Hierarchical Composite</th>
<th>Clinical Composite</th>
<th>P value Criteria</th>
<th>Single End Point–VAS AUC 5 d</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>“Trends” in 2 dyspnea end points and LOS, modest effect on WHF 5 d, large effect on death or rehospitalization</td>
<td>50</td>
<td>0.119</td>
<td>0.096</td>
<td>0.097</td>
<td>0.078</td>
<td>0.116</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.16</td>
<td>0.12</td>
<td>0.124</td>
<td>0.095</td>
<td>0.152</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.204</td>
<td>0.138</td>
<td>0.16</td>
<td>0.124</td>
<td>0.185</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.231</td>
<td>0.172</td>
<td>0.214</td>
<td>0.148</td>
<td>0.225</td>
<td>0.204</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.256</td>
<td>0.185</td>
<td>0.214</td>
<td>0.159</td>
<td>0.256</td>
<td>0.235</td>
</tr>
<tr>
<td>11</td>
<td>Very large effect on death and rehospitalization, neutral on other end points</td>
<td>50</td>
<td>0.048</td>
<td>0.063</td>
<td>0.058</td>
<td>0.052</td>
<td>0.056</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.054</td>
<td>0.091</td>
<td>0.08</td>
<td>0.068</td>
<td>0.064</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.069</td>
<td>0.096</td>
<td>0.093</td>
<td>0.084</td>
<td>0.065</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.077</td>
<td>0.136</td>
<td>0.126</td>
<td>0.108</td>
<td>0.085</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.079</td>
<td>0.138</td>
<td>0.128</td>
<td>0.111</td>
<td>0.076</td>
<td>0.063</td>
</tr>
<tr>
<td>12</td>
<td>Large+effects on all 6 end points</td>
<td>50</td>
<td>0.448</td>
<td>0.217</td>
<td>0.302</td>
<td>0.153</td>
<td>0.348</td>
<td>0.328</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.622</td>
<td>0.314</td>
<td>0.423</td>
<td>0.212</td>
<td>0.473</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.713</td>
<td>0.401</td>
<td>0.514</td>
<td>0.295</td>
<td>0.548</td>
<td>0.529</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.818</td>
<td>0.469</td>
<td>0.624</td>
<td>0.375</td>
<td>0.63</td>
<td>0.616</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.878</td>
<td>0.564</td>
<td>0.697</td>
<td>0.427</td>
<td>0.702</td>
<td>0.7</td>
</tr>
<tr>
<td>13</td>
<td>Modified large+effects on all 6 end points</td>
<td>50</td>
<td>0.4</td>
<td>0.227</td>
<td>0.259</td>
<td>0.127</td>
<td>0.317</td>
<td>0.286</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.553</td>
<td>0.327</td>
<td>0.39</td>
<td>0.181</td>
<td>0.414</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.649</td>
<td>0.408</td>
<td>0.459</td>
<td>0.228</td>
<td>0.485</td>
<td>0.473</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.762</td>
<td>0.493</td>
<td>0.561</td>
<td>0.302</td>
<td>0.57</td>
<td>0.558</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.831</td>
<td>0.573</td>
<td>0.646</td>
<td>0.354</td>
<td>0.637</td>
<td>0.627</td>
</tr>
<tr>
<td>14</td>
<td>Very large effects on all 6 end points</td>
<td>50</td>
<td>0.582</td>
<td>0.308</td>
<td>0.43</td>
<td>0.218</td>
<td>0.474</td>
<td>0.453</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.759</td>
<td>0.441</td>
<td>0.577</td>
<td>0.308</td>
<td>0.603</td>
<td>0.591</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.864</td>
<td>0.536</td>
<td>0.684</td>
<td>0.401</td>
<td>0.679</td>
<td>0.671</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.925</td>
<td>0.638</td>
<td>0.799</td>
<td>0.489</td>
<td>0.76</td>
<td>0.749</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.975</td>
<td>0.714</td>
<td>0.856</td>
<td>0.552</td>
<td>0.821</td>
<td>0.817</td>
</tr>
<tr>
<td>15</td>
<td>Neutral on all 6 end points</td>
<td>50</td>
<td>0.023</td>
<td>0.018</td>
<td>0.017</td>
<td>0.017</td>
<td>0.026</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.022</td>
<td>0.027</td>
<td>0.029</td>
<td>0.026</td>
<td>0.027</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.029</td>
<td>0.031</td>
<td>0.029</td>
<td>0.025</td>
<td>0.022</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>125</td>
<td>0.026</td>
<td>0.027</td>
<td>0.025</td>
<td>0.025</td>
<td>0.029</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>0.025</td>
<td>0.025</td>
<td>0.025</td>
<td>0.029</td>
<td>0.025</td>
<td>0.02</td>
</tr>
</tbody>
</table>

VAS indicates visual analog scale; AUC, area under the change from baseline curve; LOS, ‘WHF.’
finding can negatively affect hierarchical end points in small studies—a problem that is less likely to occur in larger studies. Interestingly, even in scenarios 9, 10, and 11, where most of the effect is observed on morbidity and mortality, the hierarchical approaches did not provide substantially more power than the average Z.

The clinical composite approach defines treatment success and failure in a more restricted way, thus many patients are left in the middle category where neither success nor failure is defined. In addition, this approach cannot really determine whether active or placebo is better if patients in both groups had events but I group lived longer than the other, or no events happened for both groups with improvement in symptoms however having different magnitudes, and so on.

With respect to analysis, nonparametric methods are appropriate for all 5 composite outcomes without assuming underlying distributions. This may be desirable with small sample sizes where distributional assumptions may not hold.

Proposed in 1984, O’Brien’s procedures for comparing treatments with multiple end points were thought very useful and various extensions of this method have been published.13,14 However, it has not been widely adapted to clinical trials today. This method is robust when the sample size is small, or the variables are not normally distributed. The null hypothesis associated with this method is no treatment effect for all end points versus the alternative that 1 treatment is better than the other for nearly all of the end points and not worse for any. Because this method uses averaged Z score across end points, its application may not be appropriate under certain situations, such as when treatment effects are expected to occur in only relatively few end points or the direction of change cannot be anticipated in advance and may not be consistent among the variables, which typically is not the case in AHF. The average Z score evaluated here treats each end point equally by averaging individual Z scores across end points. If more emphasis is desired to be placed on certain end points, the method can be modified by assigning heavier weights to certain end points. In our study, a consistently higher power was observed for this approach. The method is easy to implement and straightforward for interpretation. Thus, this approach should be considered when designing phase II trials.

At the conclusion of the phase II study, if the treatment is deemed to be of potential benefit, some of the end points may become part of the primary and secondary end points of the nascent phase III program. The end points considered for the average Z score evaluated here have been used before either alone or in combination as primary or secondary end points in similar studies.2,3,6,8,11 Advanced statistical methods could be applied to explore which end points have been most affected in phase II and should become the main end points for phase III, either as classical primary and secondary single or composite end points or comprising an average Z.

There are several limitations to the current study. First, the correlation structure used in the study was based on previous observations in trials of patients with AHF, and rounded up to be conservative for power computation. If the correlations are larger than the prespecified ones, the powers for the average Z score approach shall go down. Second, these results were not validated prospectively in a phase III AHF study, mostly because a large database for a positive AHF study where an intervention significantly affected major aspects of the AHF disease processes is not available. Hence the approach remains to be tested when a positive large phase III AHF study becomes available. Third, we chose to include composite end points focusing on positive effects of new interventions on aspects of the AHF disease process. However, the method could be extended to include whatever end points investigators feel is appropriate for evaluation of the intervention under study, and, if using the average Z approach, weighted according to relative clinical importance.

In summary, the average Z score consistently provides more power than other composite outcomes for all treatment effect scenarios examined. If large or very large effects are to be expected from a trial across 6 end points, a sample size of 100 to 150 patients per group may provide reasonable power for the study using the average Z score, approximately double the power achievable by assessing the effects on dyspnea alone.

Disclosures

G.K. serves as the principal investigator of a cooperative agreement between UNC Chapel Hill and Momentum Research Inc. Hengrui Sun, Gad Cotter, and Beth Davison are employees of Momentum Research Inc. who have provided consulting and/or trial management services to Trevena, Amgen Inc., Novartis Pharmaceuticals Corp., Novartis Pharma AG, Cardios3 Biosciences SA, Sorbent Therapeutics Inc., Celladon Corp., and Corthera.

References

Few new therapies for acute heart failure (AHF) have been approved since the introduction of furosemide in the late 1960s. This may be attributed at least in part to challenges in the design and selection of interventions and doses in phase II AHF studies. Because only a limited number of patients can be enrolled in phase II clinical studies, surrogate measures such as wedge pressure, remodeling or biomarkers have been used to identify potentially effective therapies. However, several new drugs with promising results on such surrogate measures in phase II failed to demonstrate positive effects on clinical outcomes in phase III, mostly because surrogate measures are not universally correlated with clinical benefit in AHF. Combining clinical outcomes into a composite in phase II allows simultaneous evaluation of an intervention’s effect on multiple aspects of the AHF disease process, and might allow identification of therapies that will demonstrate concordant effects on component outcomes in phase III. We have explored several methods of combining such end points. Because most clinical outcomes in AHF are only slightly correlated, such a combined end point approach increases the ability to identify interventions that potentially affect multiple clinical facets of AHF rather than surrogates. Among the composite approaches evaluated, transforming the end points to a common metric—the z score—and averaging seems to provide the highest power. The average Z approach allows a quantitative test of the overall effect on multiple outcomes, potentially allowing identification of new therapies with beneficial clinical effects in phase II that can be confirmed in phase III trials.
SUPPLEMENTARY MATERIAL:

Table 1. Power comparison at variant scenario and sample size combinations by different methods. Data were simulated with independent endpoints.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Sample Size (n per group)</th>
<th>Composite Endpoint</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average Z Score</td>
<td>Global Rank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Criterion\textsubscript{1}</td>
<td>Criterion\textsubscript{2}</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0.189</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.263</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.328</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>0.415</td>
<td>0.124</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.49</td>
<td>0.141</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>0.23</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.323</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.397</td>
<td>0.173</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>0.489</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.58</td>
<td>0.227</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>0.251</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.357</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.434</td>
<td>0.147</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>0.537</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>150</td>
<td>0.624</td>
<td>0.188</td>
<td>0.263</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>0.352</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.49</td>
<td>0.162</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.604</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>0.713</td>
<td>0.247</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.791</td>
<td>0.288</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0.408</td>
<td>0.161</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.576</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.672</td>
<td>0.303</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>0.792</td>
<td>0.353</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.855</td>
<td>0.414</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>0.207</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.286</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.353</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>0.463</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.531</td>
<td>0.124</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>0.173</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>75</td>
<td>0.251</td>
<td>0.056</td>
<td>0.069</td>
</tr>
<tr>
<td>100</td>
<td>0.313</td>
<td>0.062</td>
<td>0.099</td>
</tr>
<tr>
<td>125</td>
<td>0.403</td>
<td>0.07</td>
<td>0.092</td>
</tr>
<tr>
<td>150</td>
<td>0.466</td>
<td>0.076</td>
<td>0.116</td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>0.297</td>
<td>0.067</td>
</tr>
<tr>
<td>75</td>
<td>0.427</td>
<td>0.081</td>
<td>0.103</td>
</tr>
<tr>
<td>100</td>
<td>0.518</td>
<td>0.096</td>
<td>0.144</td>
</tr>
<tr>
<td>125</td>
<td>0.625</td>
<td>0.119</td>
<td>0.164</td>
</tr>
<tr>
<td>150</td>
<td>0.707</td>
<td>0.142</td>
<td>0.208</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>0.113</td>
<td>0.052</td>
</tr>
<tr>
<td>75</td>
<td>0.14</td>
<td>0.078</td>
<td>0.085</td>
</tr>
<tr>
<td>100</td>
<td>0.177</td>
<td>0.082</td>
<td>0.113</td>
</tr>
<tr>
<td>125</td>
<td>0.227</td>
<td>0.097</td>
<td>0.116</td>
</tr>
<tr>
<td>150</td>
<td>0.252</td>
<td>0.105</td>
<td>0.138</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>0.156</td>
<td>0.091</td>
</tr>
<tr>
<td>75</td>
<td>0.206</td>
<td>0.115</td>
<td>0.142</td>
</tr>
<tr>
<td>100</td>
<td>0.256</td>
<td>0.155</td>
<td>0.185</td>
</tr>
<tr>
<td>125</td>
<td>0.323</td>
<td>0.186</td>
<td>0.221</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>150</td>
<td>0.384</td>
<td>0.205</td>
<td>0.251</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>0.05</td>
<td>0.066</td>
</tr>
<tr>
<td>75</td>
<td>0.058</td>
<td>0.086</td>
<td>0.075</td>
</tr>
<tr>
<td>100</td>
<td>0.072</td>
<td>0.106</td>
<td>0.107</td>
</tr>
<tr>
<td>125</td>
<td>0.08</td>
<td>0.122</td>
<td>0.11</td>
</tr>
<tr>
<td>150</td>
<td>0.087</td>
<td>0.142</td>
<td>0.139</td>
</tr>
<tr>
<td>12</td>
<td>50</td>
<td>0.593</td>
<td>0.215</td>
</tr>
<tr>
<td>75</td>
<td>0.767</td>
<td>0.343</td>
<td>0.459</td>
</tr>
<tr>
<td>100</td>
<td>0.865</td>
<td>0.433</td>
<td>0.544</td>
</tr>
<tr>
<td>125</td>
<td>0.928</td>
<td>0.527</td>
<td>0.66</td>
</tr>
<tr>
<td>150</td>
<td>0.966</td>
<td>0.618</td>
<td>0.744</td>
</tr>
<tr>
<td>13</td>
<td>50</td>
<td>0.526</td>
<td>0.211</td>
</tr>
<tr>
<td>75</td>
<td>0.705</td>
<td>0.316</td>
<td>0.386</td>
</tr>
<tr>
<td>100</td>
<td>0.825</td>
<td>0.409</td>
<td>0.461</td>
</tr>
<tr>
<td>125</td>
<td>0.902</td>
<td>0.489</td>
<td>0.566</td>
</tr>
<tr>
<td>150</td>
<td>0.936</td>
<td>0.581</td>
<td>0.659</td>
</tr>
<tr>
<td>14</td>
<td>50</td>
<td>0.739</td>
<td>0.324</td>
</tr>
<tr>
<td>Flowrate</td>
<td>0.889</td>
<td>0.471</td>
<td>0.623</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>75</td>
<td>1.000</td>
<td>0.574</td>
<td>0.741</td>
</tr>
<tr>
<td>100</td>
<td>0.984</td>
<td>0.7</td>
<td>0.845</td>
</tr>
<tr>
<td>125</td>
<td>0.995</td>
<td>0.775</td>
<td>0.899</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flowrate</th>
<th>0.019</th>
<th>0.026</th>
<th>0.02</th>
<th>0.019</th>
<th>0.011</th>
<th>0.06</th>
<th>0.014</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1.000</td>
<td>0.024</td>
<td>0.018</td>
<td>0.021</td>
<td>0.06</td>
<td>0.05</td>
<td>0.018</td>
</tr>
<tr>
<td>75</td>
<td>0.025</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.008</td>
<td>0.007</td>
<td>0.027</td>
</tr>
<tr>
<td>100</td>
<td>0.025</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.008</td>
<td>0.007</td>
<td>0.027</td>
</tr>
<tr>
<td>125</td>
<td>0.022</td>
<td>0.029</td>
<td>0.023</td>
<td>0.033</td>
<td>0.007</td>
<td>0.004</td>
<td>0.021</td>
</tr>
<tr>
<td>150</td>
<td>0.02</td>
<td>0.024</td>
<td>0.028</td>
<td>0.022</td>
<td>0.008</td>
<td>0.006</td>
<td>0.017</td>
</tr>
</tbody>
</table>