Introduction
Heart failure (HF) is a leading cause of hospitalization. There are few tools to accurately identify patients at high risk for unplanned admission in the outpatient setting. We used machine learning (ML) on outpatient electronic medical records and medical claims to develop a HF specific predictive model.

Methods
The OM1® Cardiology data warehouse contains deep clinical and claims data on patients seen in cardiology practices across the US. HF Patients with at least 18 months of data (July 2013 to December 2016) were included, with the last 6 months serving as the prediction period. The outcome was unplanned admission due to HF during the prediction period.

Results
A total 48,761 patients with 2,770 unplanned admissions were included in the analysis; median age was 71 years, 49% were men, and 54% were white (Table 1). The top 5 predictors from the training set, as determined by machine learning, were OM1 medical burden index (ML derived score on a 0-1000 scale, leveraging all available prior history), number of services in the past 12 months, number of hospitalizations in the past 12 months, time since last hospitalization, and Charlson comorbidity. These 5 predictors were used to derive HF risk scores in the validation set of 14,546 patients (Table 2). Using the OM1 medical burden index as the single predictor, the model correctly predicted outcomes for 12,189 (84%) patients. All top 5 predictors identified by machine learning were statistically significant in the regression model (Figure 2). The odds ratios in Figure 2 represent one unit increase of OM1 medical burden index on a log scale, every 20 units increase of OM1 medical burden index on a log scale, every 5 units increase of OM1 medical burden index on a log scale.

Table 1. Patient characteristics during the 12-month observation period, stratified by the admission status during the 6-month prediction period

<table>
<thead>
<tr>
<th>Variable</th>
<th>Admitted (N=2,770)</th>
<th>No admission (N=23,385)</th>
<th>Total (N=26,155)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>Female (n=1,329)</td>
<td>Male (n=1,441)</td>
<td></td>
</tr>
<tr>
<td>Median Age (years)</td>
<td>70 (Q1-Q3: 62-79)</td>
<td>71 (Q1-Q3: 69-78)</td>
<td>71 (Q1-Q3: 62-79)</td>
</tr>
<tr>
<td><strong>Clinical Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>1.6 (Q1-Q3: 0-3)</td>
<td>1.9 (Q1-Q3: 1-4)</td>
<td>1.7 (Q1-Q3: 1-3)</td>
</tr>
<tr>
<td>OM1 medical burden index</td>
<td>387 (Q1-Q3: 150-940)</td>
<td>432 (Q1-Q3: 130-1060)</td>
<td>409 (Q1-Q3: 130-1060)</td>
</tr>
<tr>
<td><strong>Patient Encounter Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time since last hospitalization</td>
<td>11 (Q1-Q3: 1-4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of hospitalizations</td>
<td>2.9 (Q1-Q3: 1-6)</td>
<td>1.6 (Q1-Q3: 1-4)</td>
<td>2.2 (Q1-Q3: 1-4)</td>
</tr>
<tr>
<td>Time since last hospitalization</td>
<td>11 (Q1-Q3: 1-4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>1.6 (Q1-Q3: 0-3)</td>
<td>1.9 (Q1-Q3: 1-4)</td>
<td>1.7 (Q1-Q3: 1-3)</td>
</tr>
</tbody>
</table>

Table 2. The OM1 medical burden index is the single most predictive variable in univariate analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Omn1 Medical Burden Index</th>
<th>Odds Ratio 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C statistic</td>
<td>0.72</td>
<td>1.04-1.09</td>
</tr>
<tr>
<td>Number of services in the past year</td>
<td>0.71</td>
<td>1.01-1.06</td>
</tr>
<tr>
<td>Time since last hospitalization</td>
<td>0.70</td>
<td>1.01-1.06</td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>0.66</td>
<td>1.08-1.15</td>
</tr>
</tbody>
</table>

Conclusions
We demonstrated the utility of machine learning in leveraging variables readily available in an outpatient EMR and medical claims to predict hospitalizations in 8 out of 10 patients (C Statistic: 0.79). When integrated into the clinical workflow, such tools may offer the ability to focus resources on patients at highest risk for unplanned admission.