The Hilltop Institute

The Hilltop Pre- Models Predictive Power
April 2023

The Hilltop Pre- Models use information on patients' utilization, medications, comorbidities, social determinants of health, and other risk factors to generate monthly risk scores. These risk scores, transmitted to providers via CRISP (Chesapeake Regional Information System for our Patients), are intended to capture beneficiaries' risk of experiencing specific outcome events in the near future and rank attributed beneficiaries in each practice's panel based on their risk of experiencing the model-specific event. For each Medicare beneficiary who receives primary care through Maryland Primary Care Program (MDPCP) participant practices, Hilltop estimates three event risk scores and accompanying reasons for risk using the following models[1]:

- **The Pre-AH Model™** predicts risk for an inpatient hospitalization or emergency department (ED) visit for certain conditions that are believed to be preventable through the receipt of timely, quality primary care. This model generates the "Avoidable Hospitalizations (Pre-AH)" scores, which were first released on October 11, 2019, and have been updated monthly since.

- **The Pre-DC Model™** predicts risk for an inpatient hospitalization or ED visit due to *severe* complications of type 2 diabetes across six domains of complications: ophthalmic/retinopathy, nephropathy, cerebrovascular, cardiovascular, peripheral vascular, and metabolic. This model generates the "Severe Diabetes Complications (Pre-DC)" scores, which were first released on October 14, 2022.

- **The Pre-HE Model™** predicts risk **of eligibility for hospice** in order to identify patients who are potentially appropriate candidates for advanced care planning and provide care teams with information that can guide the sensitive and difficult conversations about end-of-life care with patients and their families. This model generates the "**Hospice Eligibility and Advanced Care Planning** (Pre-HE)" scores, which were first released on October 14, 2022.

Each risk score is designed to assist providers in proactively identifying patients at a high risk of incurring the model-specific outcome. Using this information in conjunction with clinical judgement, providers can make informed decisions about allocating scarce care coordination resources, directing these resources to the individuals who will benefit from them the most. Given this goal, it is crucial that the Pre- Model risk score rankings are accurate. Individuals with a high likelihood of incurring the model-specific event should receive high risk scores, and individuals with a low likelihood should receive low risk scores.

---

[1] For additional information about each model, see *Risk Score Specifications and Codebook for The Hilltop Institute's Pre-Models Version 1* on the MDPCP "Practice Information" [webpage](#).

The Hilltop Institute designed each of the Pre- Models to prospectively maximize the accuracy of the risk score ranking. To do this, we feed into the model a pool of approximately 200 risk factors that have already been demonstrated, in published research, to be predictive of the model specific outcome (e.g., avoidable hospitalizations). We let the data speak for itself in selecting which risk factors are most important. We use a split-sample development process to ensure that we are not overfitting the model. Finally, we use a shadow coding system to ensure that we are not introducing error into the model that would weaken the predictions. Good model design, however, is no guarantee of good predictive power.

Model Predictive Performance

One of the ways Hilltop gauges the predictive accuracy of the Pre- Models is to look at how well the risk score rankings for a given month (for example, November 2021) correspond to who *actually* experiences the outcome event in the following month (for example, December 2021) using new data. By using this approach, we can determine that, for example, 50% of the riskiest individuals as ranked by the Pre-DC Model™ in November 2021 incurred a diabetes complication event in the following month. This example would be suggestive of good model performance: the individuals that our model predicts to be very risky do, in fact, have a high risk of incurring the model specific event. We prefer to evaluate predictive accuracy post-deployment using production data from the month following release of the risk scores. However, when future production data are not yet available, as is the case with the new Pre-HE and Pre-DC Models, we can evaluate predictive accuracy during the model development process using holdout data that are not used to build the model. It is important to note that, in either approach, the model has no knowledge of who, in the following month, would experience the outcome event; that is, the scores are purely predictive.
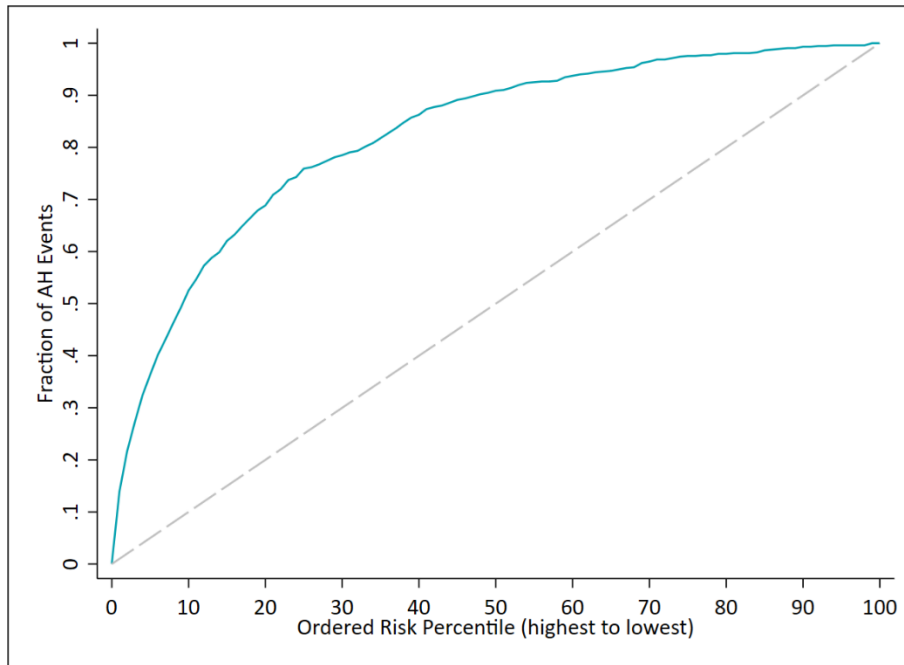
Due to the nature of the modeling problem—estimating the distribution of risk, rather than binary classification—we capture program-wide model performance using the *concentration curve*. This allows us to answer the question, "Of all the people who have a model-specific outcome event (e.g., an avoidable hospital event) in a given month, what percentage is accounted for by the top 10% riskiest individuals as ranked by the associated Hilltop Pre- Model (e.g., Pre-AH Model™)?" Intuitively, if the answer is close to 100%, then the model is performing very well: all the people who have avoidable hospital events are captured by high risk scores. Analogously, if the answer is close to 10%, then the model has no predictive value and is essentially ranking people at random.

Hilltop assesses predictive performance separately for each model (see Table 1 for summary).

Table 1. Summary of Predictive Performance by Model

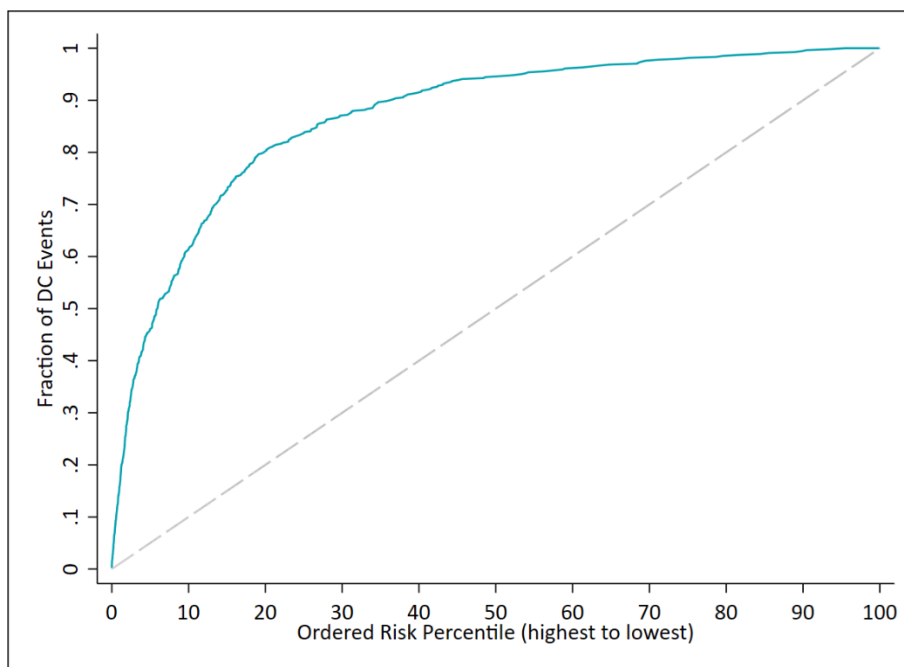|	Pre-AH	Pre-DC	Pre-HE
Percentage of Outcomes in Top 10% of Risk Scores	47%	61%	63%

The Hilltop Institute

Pre-AH Model™

Risk Score Rankings from: April 2022

Event Outcomes from: May 2022

Evaluation Data: Production data from the month following score release

Summary: Approximately 47% of all individuals experiencing an avoidable hospital event in the following month are contained in the top 10% riskiest individuals as ranked by the Pre-AH Model™. We interpret this as good performance of the model, and the other months show similar results (range: 47.8%-57.7%). These results imply that, if care managers were to focus their efforts on the top 10% riskiest beneficiaries each month, then they could reach almost half of all individuals experiencing avoidable hospital events that month. Additionally, this measure does not account for the individuals that *would* have incurred an avoidable hospital event but did not due to receipt of advanced primary care; therefore, the true accuracy of the Pre-AH Model™ risk scores may be even higher that what is reported here.

The Hilltop Institute
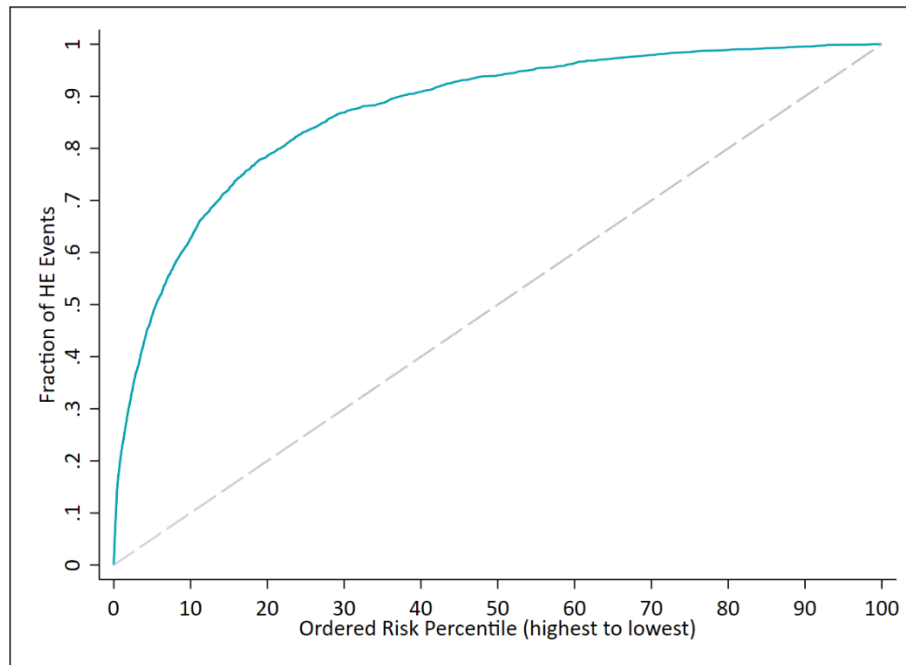
Pre-DC Model™

Risk Score Rankings from: April 2022

Event Outcomes from: May 2022

Evaluation Data: Holdout data during model development

Summary: Approximately 61% of all individuals experiencing a severe type 2 diabetes complication event in the following month are contained in the top 10% riskiest individuals as ranked by the Pre-DC Model™. We interpret this as good performance of the model, and the other months show similar results (range: 58%-63%). These results imply that, if care managers were to focus their efforts on the top 10% riskiest beneficiaries each month, then they could reach more than 60% of all individuals experiencing diabetes complication events that month.

The Hilltop Institute

Pre-HE Model™

Risk Score Rankings from: October 2021

Event Outcomes from: November 2021

Evaluation Data: Holdout data during model development

Summary: Almost 63% of all individuals with outcome events are contained in the top 10% riskiest individuals as ranked by the Pre-HE Model™. We interpret this as good performance of the model, and the other months show similar results (range: 61.9%-64.6%). These results imply that, if care managers were to focus their efforts on the top 10% riskiest beneficiaries each month, then they could reach more than 60% of the patients who may be appropriate candidates for hospice care to proactively begin advanced care discussions.

Note: due to the timing and availability of death date data, we evaluate predictive accuracy for this model for an earlier period than the other models.

The Hilltop Institute