Quantitative Medicine Project Meeting with Lundbeck
Wednesday October 9th, 2013
Webinar

Cambridge Healthtech Associates	**Quantitative Medicine**
Ernie Bush Dawn Van Dam Amid Zand **Lundbeck** Ulf Norinder Jorrit Hornberg	John King Scott Bodine Geoff Hoare Josh Kangas David Demosthenes *Jim Parrino (not on call)* *Bob Murphy (not on call)* *Jamie Grooms (not on call)*

I) Welcome

Dawn started the meeting with introductions and Josh presented the results of Quantitative Medicine's analysis on the hepatotoxicity study design and results.

II) Quantitative Medicine Presentation

Josh unveiled the new acronym for their proprietary computation system. The system is now called Computational Research Engine (*CoRE*™).

Josh discussed the agenda of this meeting:

1. Hepatotoxicity Study Design
Josh described the problem statement and the approach he took to conduct the study.

As he explained, the problem was to accurately predict toxicity for new compounds based on the results of the High Content Screening (HCS) experiments. The approach

was defined by testing each new compound in the HCS assay and using the results as inputs for a predictive model learned from prior experimentation.

2. Simulation on Results

The Lundbeck data included 105 compounds and 8 HCS measurements. There was an issue with some of the data provided by Lundbeck, therefore those data points were not included in the study.

Josh discussed the simulation protocol based on the Active Learning approach. The dataset was randomly divided into a training set (80%) and a test set (20%). All results were hidden from the *CoRE™* and for each round a compound was selected from the training set and the toxicity and HCS readouts were revealed to the learner. Finally, a predictive model was learned from only revealed training data and predictions were made for all compounds in the testing set.

Ulf: With how many training set compounds did you start with for the first round?

Josh: I started with 2 compounds in the training set, one toxic and one non-toxic compound.

Josh described the receiver operating characteristics. He explained that the final prediction was defined as the hepatotoxicity for all 'unobserved' experiments in the testing set and the accuracy of those predictions were measured by calculating an ROC curve. Severe and withdrawn cases were all considered toxic for this study.

Josh then discussed the comparison methods. For this study, he defined the standard approach as random selection of the training set followed by a RandomForest Prediction Model. This method was compared with Quantitative Medicine's Computational Research Engine. In the *CoRE™* approach, a single compound was selected for execution and a model was learned using prediction methods based on all revealed HCS results and QM's library to predict toxicity.

Josh shared the results of the project. He shared a graph of a number of compounds from the training set (x-axis) vs. area under the ROC curve (y-axis). The random baseline showed an area under the ROC curve of ~0.5.

The RandomForest and *CoRE™* methods were compared next. RandomForest showed a spike around 10 compounds, which was determined statistically insignificant. It was also shown that the predictive accuracy of the model learned using the *CoRE™* is in expectation higher than standard practices. Josh explained that this is due to the additional learning provided by the *CoRE™* and the way data are analyzed. He added that RandomForest is only one of the machine learning methods used in the *CoRE™* to analyze different data sets.

Ulf: As you increase the number of training set compounds, is it done systematically or in random order?

Josh: RandomForest does each addition in random but our method does that systematically to ensure that each compound adds additional data.

Josh demonstrated how *CoRE*™ substantially improves the accuracy of predictions in comparison to RandomForest.

Ulf: What would happen if you didn't add the extra information? How does your machine learning system perform without adding extra information?

Josh: I can get back to you on that question.

Ulf: How do you do the linkage between the neighboring compounds?

Josh: The external information used in this study was from identical compounds. For 20% of the compounds in your database, we only had structural information. Therefore, the neighboring compound information was not used for those compounds.

Josh then compared the results of this study to those of a ToxCast simulation study. He showed the results graphically (x-axis: Percentage of experimental space explored; y-axis: area under ROC curve). He showed that about 80-90% of the space had to be studied for standard methods, whereas the *CoRE*™ only needed to evaluate 10-20% of the space.

When the ToxCast study and the Lundbeck study were compared, there was a difference in the cross-over point and a difference in the fraction of experiments required to reach maximum accuracy. In addition, the accuracy of the Lundbeck study improved continuously; whereas the ToxCast study accuracy improvement stopped after 40% of the experimental space had been explored. Josh argued that this is due to the quality of data, selection of compounds, and dimensionality of the experimental space.

Ulf mentioned that the HCS measurement data needs to be removed and not be used in the analysis since this data is not available for compounds that have not yet been synthesized. There would also be no identical information available; only neighboring information.

3. **Prospective Uses of the Established *CoRE*™ Model**
   Josh described how studies could be designed for single compounds, multiple compounds, or multiple assays (prioritization).

4. **Next Studies**
   Ulf suggested that a head-to-head comparison between RandomForest and *CoRE*™ without the extra data could be a good option for the next study.

   Josh agreed to run this study and also suggested two additional studies.

Study 1: Neighboring data will be used, but actual experimental results will not be used.
Study 2: Look into how many assays can be removed without reducing accuracy.

IV) Action Items

1) Webinar and recording will be sent to Lundbeck.
2) Three studies discussed will be evaluated in the next meeting.
