Accelerating the Identification of New and Unexplored Immuno-Oncology Therapies with IBM Watson

Author: John Gregory, Director, Pfizer R&D Project & Portfolio Management Business Technology
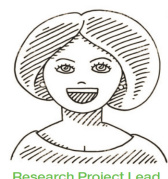
The Business Challenge:

" As an Oncology R&D scientist, it is currently time consuming and difficult for me to review the vast sea of research publications and data in order to identify, and rank, promising novel immuno therapies, and predict the viability of those therapies in terms of efficacy and toxicity "

Research Project Lead

The Result

Early use of Watson Drug Discovery for I-O has helped our Oncology R&D scientists to:

✓ Identify new tumor-types that have a similar representation of genes being modulated to help predict new indications

✓ Validate current sets of gene targets we are pursuing for immune response

✓ Identify novel gene sets that have not been associated with immune response that fight the tumor in a different way
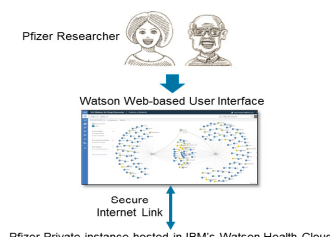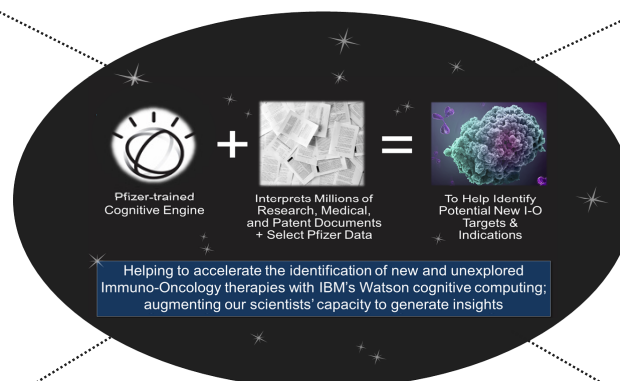
The Opportunity:

How can we leverage IBM Watson Drug Discovery (WDD) to review millions of research papers, applying Natural Language Processing, Machine Learning & Predictive algorithms to help our scientists accelerate the identification of new and unexplored Immuno-Oncology therapies with cognitive computing?

Watson Health™

Pfizer-trained Cognitive Engine + Interprets Millions of Research, Medical, and Patent Documents + Select Pfizer Data = To Help Identify Potential New I-O Targets & Indications

Helping to accelerate the identification of new and unexplored Immuno-Oncology therapies with IBM's Watson cognitive computing; augmenting our scientists' capacity to generate insights

Pfizer Researcher

Watson Web-based User Interface

Secure Internet Link

Pfizer Private instance hosted in IBM's Watson Health Cloud

Corpus
Includes: ~26M research abstracts, 1.5M full-text medical journal articles, 4M patents, and select Pfizer data

How Watson Applies its Cognitive Engine:

Understands…
Watson can read & understand documents & data (both structured and unstructured) at a massive scale

Reasons…
Watson searches & analyzes data, returning evidence-based rankings and predictions

Learns…
Watson learns and adapts over time through interacting with the data and users

WDD aggregates information from a large number of sources, including:
- Medline abstracts
- Full-text journals & medical textbooks
- Chemical databases
- Patents
- Ontologies & Dictionaries (gene cell, drug, AE, etc.)
- Pfizer pre-clinical tox reports

Lessons To-Date: The Most Important Consideration for a Machine Learning Platform is the Quality of Inputs:

- **Importance of Annotation (training) Process**: Significant effort to train Watson through annotating ~150 research papers over a 3 month period

- **Implications of using Licensed Content from external vendors**. Even data sources that are "public" may have usage terms associated with them for commercial entities that Pfizer has or needs to negotiate. Important to engage Legal and Information & Library Sciences early

The Implementation

Two primary use cases were used to configure & train Watson:

1. Predicting Combination Therapy Efficacy — Can Watson help our scientists to **identify effective drug targets (genes) for specific tumor types** that might have otherwise been missed?

2. Predicting Toxicity — Can Watson help Pfizer **identify potential safety issues,** accelerating decisions to pursue other targets or mechanisms of action?

How We Trained Watson – The Annotation Process

Step 1: Pfizer Research Scientists identified a set of ~150 research papers that are representative of a broad spectrum of scientific research in the field of Immuno-Oncology
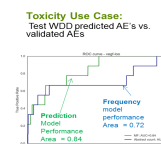
Step 2: Pfizer Research Scientist used IBM's 'Watson Knowledge Studio' tool to show Watson how to interpret the text by annotating each research paper – identifying/highlighting the entities (e.g. drugs, genes, diseases) and relationships between those entities

"…doxorubicin results in extracellular signal-regulated kinase (ERK)2 activation, which in turn phosphorylates p53 on a previously uncharacterized site, Thr55…"

How We Validated Watson's Predictive Strength

Purpose: Test that the predictions presented by WDD match with the Pfizer expert and/or trusted external data for specific use cases, and benchmark this against a background frequency ranking

- Each use case used for validation included the **input** (aka training set) the user would give WDD, and the **output** the user expects in WDD (a.k.a. test/validation entities)

- We tested how well we could predict the validation set

- We compared (benchmark) WDD's ranking to a background frequency ranking

- We measured Watson's model performance with standard methods such as the ROC (Receiver Operator characteristics) curve

Toxicity Use Case: Test WDD predicted AE's vs. validated AEs

Prediction Model Performance Area = 0.84

Frequency model performance Area = 0.72

Outcome:
The Area under the curve for the ROC plot is greater for the prediction model than the model based on frequency