Effectively predict patient discontinuation with machine learning/AI and opportunities for Rx switching using real time competitive market alerts

Ryan Sun, Sr. Director of Analytics at Claritas Rx

Presenting Author

Ryan Sun

Sr. Director of Analytics

Claritas Rx

Agenda

- Patient adherence challenge impacting commercial teams within biopharma today

- Diverse data sources available in Claritas data ecosystem

- Apply machine learning to predict patient discontinuation risk to support timely patient intervention

- Opportunity for Rx switching using competitive market alerts

Patient access and adherence – a challenge for manufacturers

Data Quality & Accessibility
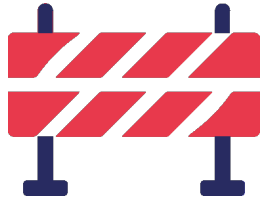
More Payer Controls
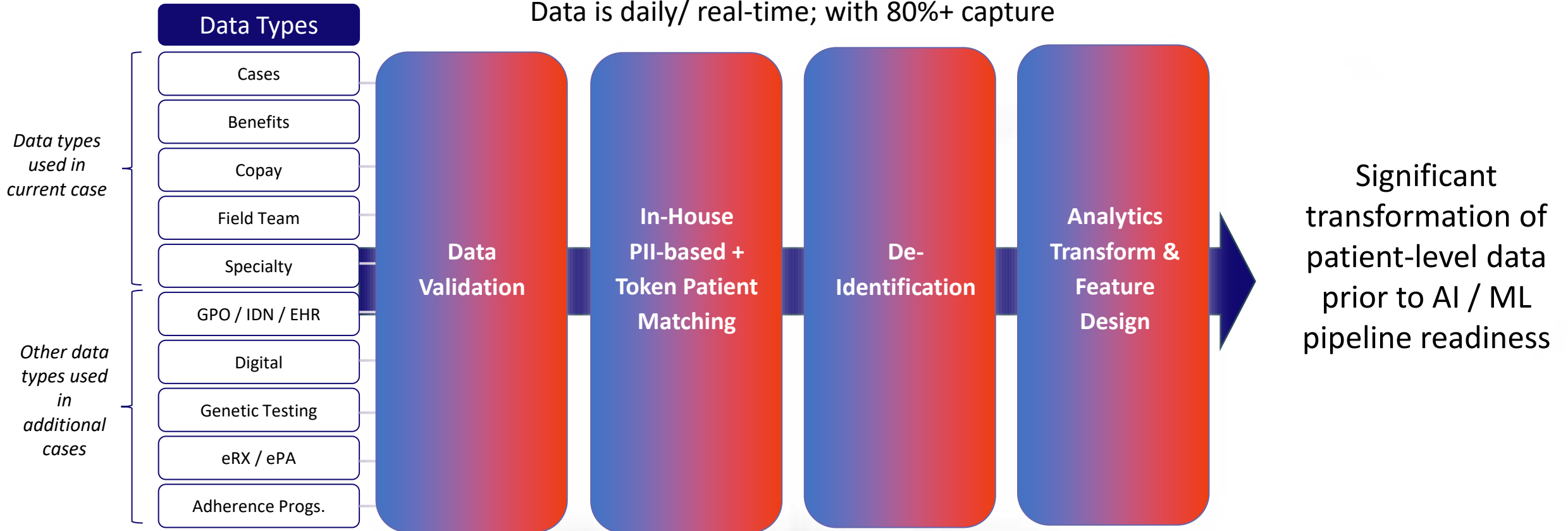
Adverse effects

Higher Patient Abandonment

Objective: apply machine learning to predict which patients are likely to discontinue treatment

Claritas Rx Data Ecosystem

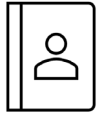Scalable solution for integrating fragmented data for diverse users

Data is daily/ real-time; with 80%+ capture

Data Types

Data types used in current case
- Cases
- Benefits
- Copay
- Field Team
- Specialty

Other data types used in additional cases
- GPO / IDN / EHR
- Digital
- Genetic Testing
- eRX / ePA
- Adherence Progs.

Data Validation → **In-House PII-based + Token Patient Matching** → **De-Identification** → **Analytics Transform & Feature Design** → Significant transformation of patient-level data prior to AI / ML pipeline readiness

200+features as predictive variables

Developed across the patient treatment journey

- Optional hub, specialty medicine

- No single dominant predictor; many (weak) variables **in combination** needed to make a stronger model

Demographics	Prescriber	Payer	Treatment	Patient Support Service
Age	Physician Specialty	Payer Type	First Time vs Restart	PAP/ FREE Drug Utilization
Gender	Academic vs. Community	PBM	Time lapsed since last shipment	Patient Support Call Frequency
Dx and Comorbidities	Prior patient adherence	Out of Pocket Cost	Titration Pattern	Patient Education Program
Prior Medications	Physician Experience with Product (# of Patients)	Prior Auth / Appeal Status	Prior Adherence	Patient Interaction Notes (Free Text)
…	…	…	…	…

Extracting Insights from Patient Interaction Notes

Compliant Natural Language Processing (NLP)

Unstructured notes from providers contain significant predictive value when parsed

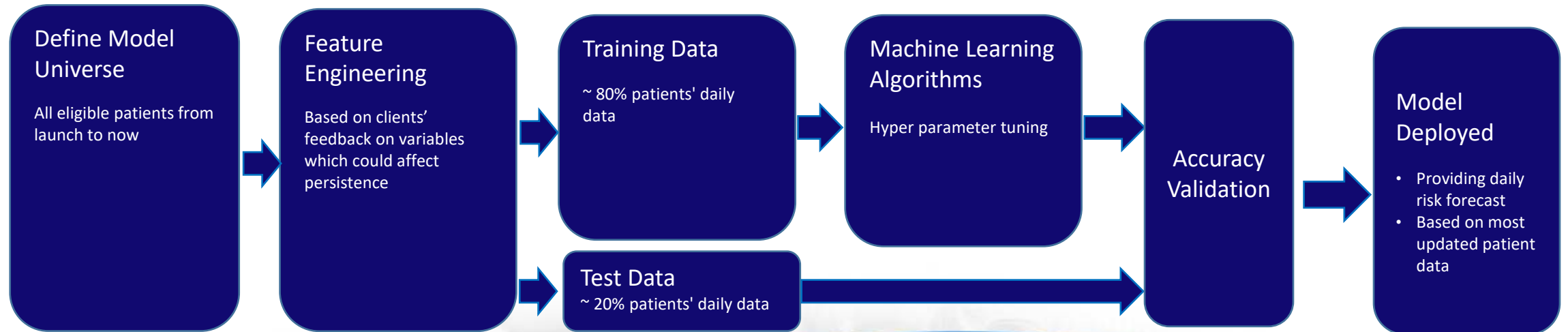Extract correlation factors from patient engagement notes:

✓ Clinical symptom identification

✓ Patient sentiment quantification

✓ Text vectorization

*Receiving and analyzing these identifiable data may have privacy considerations

Model Development Methodology

- Claritas Rx has explored a variety of ML model types for this problem
  - ✓ (Good old-fashioned) logistic regression
  - ✓ Artificial neural network
  - ✓ Support vector machine
  - ✓ Random forest
  - ✓ Gradient boosting decision tree (XGboost, Catboost, etc)

- The winning algorithm type: Gradient Boosting Decision Tree

Define Model Universe	Feature Engineering	Training Data	Machine Learning Algorithms	Accuracy Validation	Model Deployed
All eligible patients from launch to now	Based on clients' feedback on variables which could affect persistence	~ 80% patients' daily data	Hyper parameter tuning		• Providing daily risk forecast • Based on most updated patient data

Test Data
~ 20% patients' daily data

Wide Net vs Targeted Approach to Interventions

- Model performance evaluation at patient level
  - Precision: TP / [TP + FP] (% of model's positive predictions turn out to be true)
  - Recall: TP / [TP + FN] (% of actual positives are predicted by the model)

- Manufacturer can decide on a probability threshold appropriate for your resources.
  - WIDE NET approach: Manufacturers with extensive patient support staff may decide to set a lower threshold of 30% (meaning any patient with a 30% or greater risk of DC will be evaluated by the manufacturer staff)
  - TARGETED approach: Manufacturers with limited staff may want to set a higher threshold of perhaps 70%

Prob. Threshold	Risk Category	Test Patient Count	Actual DC	Remain Active	True Positive	False Positive	True Negative	False Negative	Precision	Recall
70%	Very High	2,465	690	1,775	359	58	1,716	331	86%	52%
50%	High	2,465	690	1,775	483	170	1,605	207	74%	70%
30%	Medium	2,465	690	1,775	580	580	1,195	110	50%	84%

* Based on model evaluation metrics synthesized across multiple customers to protect client privacy

Positive Customer and Patient Impact

Case Study: Patient would have fallen through the cracks without our predictive model

Background

- (2021) Implemented a discontinuation prediction model for patients on a rare disease therapy, to help the manufacturer identify patients at risk.
- (Q2' 22) Patient advocate user of our model found a long-time patient that had moved into an escalated risk of discontinuation.
- Advocate did not have a relationship with the patient directly, but notified the sales rep that they may want to speak with the patient's doctor (the patient had signed a HIPAA waiver).

Approach

- Sales rep spoke to the doctor who had no idea that the patient stopped their medication.
- Doctor learned that while on vacation, the patient had been taken off drug by an inexperienced HCP due to a minor tolerability issue noted by the patient.

Result

- The doctor called the patient and restarted on treatment right away, resuming a therapy that had provided long term control of the patient's disease.

- From the patient advocate's point of view, this patient would have fallen through the cracks without our predictive model

New, Unique Claritas Rx Market Data

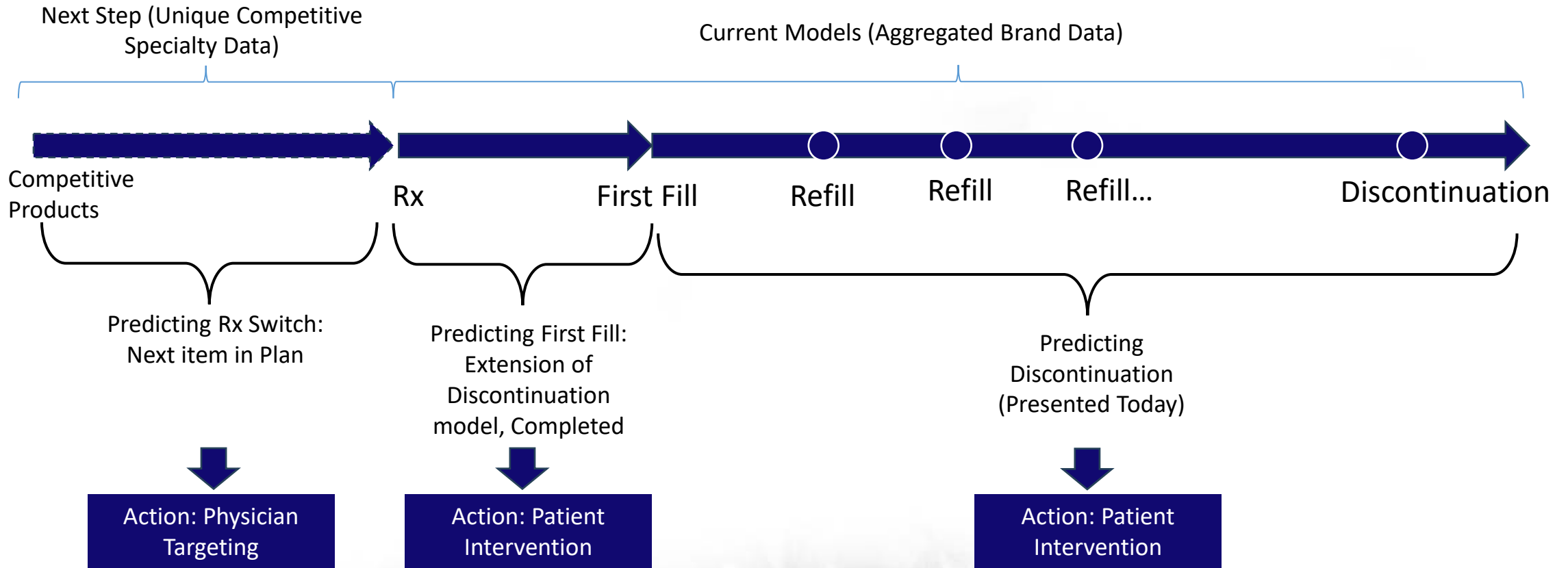Opportunity to apply AI / ML to new data sets across brands

Significantly larger patient volumes available for predictive analytics than brand case

Novel Claritas Rx Data Set	
Daily Data Across Specialty Brands	• Traditional Rx claims have 2-8 weeks of lag from patient event to analytics availability date • Claritas Rx data is on a 1–5-day lag
Visibility From Rx to Discontinuation	• Visibility into pre-dispense milestones • Visibility into Rx abandonment and treatment discontinuation dates and reasons • payer controls, out-of-pocket, and other access barriers
Can be Integrated with Claims for Max Breadth	• Significant (~95%) overlap between Claritas Rx specialty data and best-practice Rx claims

Objective: predict daily when patients will drop off *competitor* brands using novel specialty market data set

pmsa
PHARMACEUTICAL MANAGEMENT
SCIENCE ASSOCIATION

Predicting Rx Switch

Increase the "new patient funnel" by using our daily specialty data

Opportunities to apply AI/ML to new data sets across brands

- **Predict patients likely to discontinue** - AI / ML can be applied to open hub specialty brands, using many weak variables in combination to predict patients likely to discontinue

- **Enabling a broader set of data vs. typical models** - Data integration, quality, and privacy methods for structured and unstructured data enable a broader set of data to be leveraged than may be typical

- **It's only the beginning-** future opportunities exist to leverage new, daily specialty data across the market to better predict switch Rx opportunities