lumiata

Making AI work in healthcare

How GPU-accelerated AI can help us predict chronic disease amongst billions

Ash Damle
Founder & CEO
ash@lumiata.com
@ashdamle

Data Rich

Insight Starved

Resource Strapped

∴

Care Poor

Mary, Age: 67
BMI: 27.82
with

> 530 data points

12+ Conditions

angioedema, benign paroxysmal positional vertigo, depression, diabetes type 2, dyslipoproteinemia, fungal infection nails, hypercholesterolemia, hyperglycemia, hypertension, hypokalemia, hypotension, intertrigo, LV hypertrophy, major depressive disorder, mitral regurgitation, mitral valve prolapse, orthostatic hypotension, peripheral vascular disease, ...

26 Labs
4 Meds
28 Visits
1 Admissions
1 30 day readmit
$10K+ Paid Out

Imagine a world where health data is put to work

everyday, every-minute, everywhere

Imagine a world of perfect health risk awareness

Forestall and avoid preventable disease. Save billions of dollars. Ensure longer, happier lives.

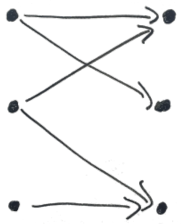
But our health is complex: 37+ trillion cells & counting

With labs, procedures, meds, diagnoses, time and more, there are millions of different variables per person.

Data Challenge 1 | Health Data is Dirty, Incomplete and Fuzzy

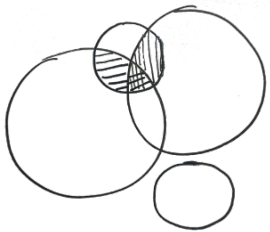
Missing Data

No Lab Units or Ranges



N-M Mappings

N ICD 9 → M ICD 10



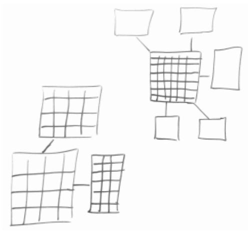
Fuzzy & Overlapping Classifications

NDC → RX

Inconsistent Data

Clinical Notes = Unstructured

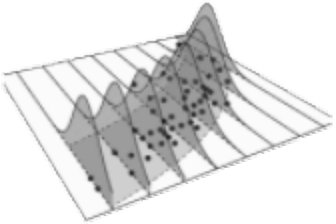
Data Challenge 2 | It's Sparse and Fragmented



Fragmented Records

Average of 3.5 different data sources for same patient

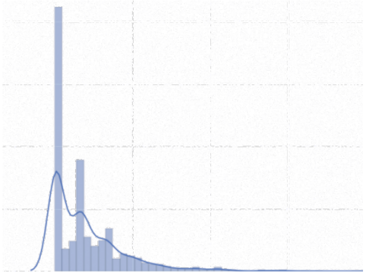
Many times PCP knows about less than 30% of patient data



Infrequent & Stochastic Sampling

Labs and other variables are not checked each time

No medical info when patient is well



Missing Key Data & 3 Year Churn

> 20% of patients appear to be submarine

Data Challenge 3 | And Has Super High Dimensionality

1.1M	600k	4.5M	2.5M	200K	2.5M
Condition Features	Procedure Features	Medication Features	Lab/Imaging Features	Provider Features	Unstructured/ Other Features

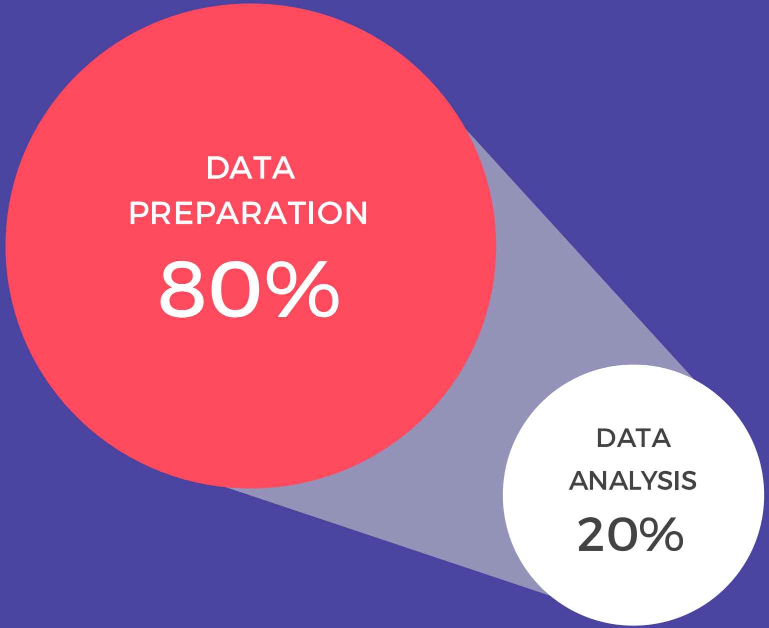
With labs, procedures, meds, diagnoses, and more combined with temporal patterns, there are millions of different potential variables per person

Healthcare Data's Huge Opportunity is Unrealized

- ## Data
  Messy, incomplete, and fuzzy
  Sparse, fragmented, and difficult to combine
  Super high dimensionality

- ## Insight
  Low precision

- ## Engagement
  Weak clinical reasoning for follow-up
  Sub-optimal chase lists with very low ROI



https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/#3325b1f66f63

To prevent and forestall chronic disease,

we need innovations to manage the complexity of health data so we can make the most of it.

So why Healthcare AI now? GPUs make it computationally tractable.

- Speed: 100x speed up makes iteration and experimentation feasible
- Precision: healthcare needs high precision and Deep Learning enables a significant boost in performance in high dimensional spaces
- Transparency: Deep Learning models that interpret Deep Learning Models requires 10x+ the computation
- Prescriptive & Predictive: optimization simulations on top of predictions require 10x+ the computation

		Desktop CPU				Server CPU						GPU		
		1	2	4	8	1	2	4	8	16	32	G.980	G.1080	T.K80
ResNet	Caffe	7.810	5.312	**4.056**	5.876	6.150	5.390	4.314	**4.124**	4.500	5.034	-	**0.208**	0.353
	CNTK	-	-	-	-	-	-	-	-	-	-	0.289	**0.261**	0.468
	TF	21.63	12.19	7.655	**6.340**	20.49	14.340	7.703	4.600	**2.890**	3.937	0.226	**0.085**	0.392
	Torch	12.10	7.147	**-**	-	10.16	6.928	4.856	3.757	**3.524**	4.165	0.216	**0.181**	0.412
LSTM-32	CNTK	0.579	0.391	**0.306**	1.153	0.591	0.418	0.353	0.338	**0.342**	0.442	0.433	**0.366**	0.602
	TF	9.305	3.432	**2.020**	1.722	6.453	3.782	2.167	1.228	**0.769**	0.706	0.086	**0.083**	0.122
	Torch	4.872	2.680	**2.366**	3.645	4.704	2.971	2.067	1.706	**1.763**	2.900	0.124	**0.098**	0.204
LSTM-64	CNTK	1.026	0.690	**0.535**	1.860	1.043	0.756	0.622	0.585	**0.648**	0.790	0.779	**0.649**	1.052
	TF	11.69	7.292	3.515	**3.476**	12.76	7.823	4.402	2.524	1.590	**1.469**	0.178	**0.173**	0.233
	Torch	9.622	5.323	**4.980**	6.975	9.364	5.613	4.054	**3.252**	3.357	5.815	0.247	**0.194**	0.406

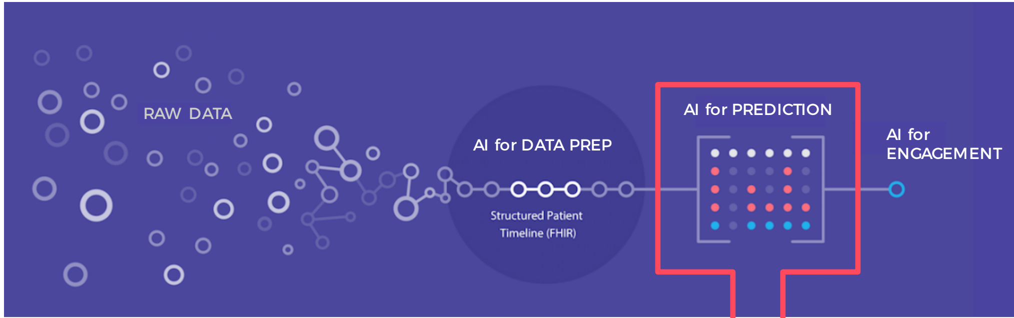
https://www.nextplatform.com/2016/09/01/cpu-gpu-put-deep-learning-framework-test/

NVIDIA lumiata

Powering Artificial Intelligence for Healthcare through GPUs

- Increased processing speed
- Reduced infrastructure complexity
- Increased model accuracy
- More precise predictions on individual health

LUMIATA AI | The Intersection of AI and the Prediction of Chronic Disease



Connecting dots to drive earlier, more accurate provider and patient engagement

1. How a person's health is likely to change
2. When the change may occur
3. What supporting clinical factors to evaluate

LUMIATA AI | Built on Growing Data Assets, Achieving More Perfect Predictions

A Health "Brain" Getting Smarter Everyday With Deep Learning + Medical Science

We've been busy training that brain in very specific ways – sending it to MD, MBA, and Actuarial school, if you will...

175M+
Patient record years

40M+
Connections between medical concepts

39K+
Physician curation hours

3TB+
Unstructured data

50M+
Articles mined from PubMed

60M+
Patients

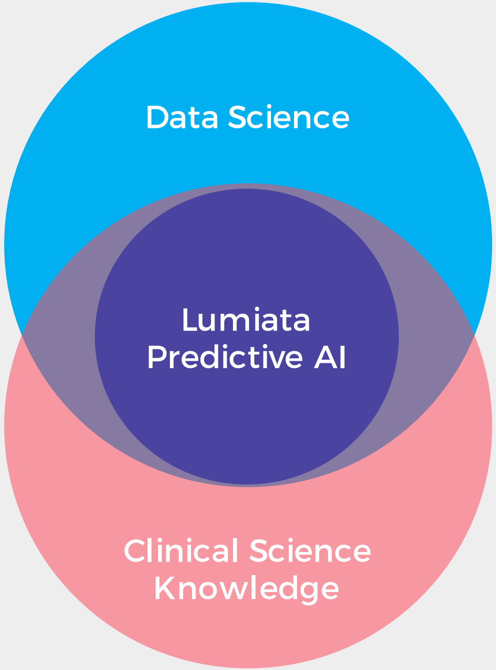
Built on Massive Foundational Data Sets and Sources

LUMIATA AI | Powered By a Deep-Learning-Based AI Stack Focused on Healthcare

RAW DATA	1 AI for DATA PREP	2 AI for PREDICTION	3 AI for ENGAGEMENT

	10x	30%+	3x+
	faster data processing with standardized data representation across multiple data sources	more accurate clinical predictions	ROI from Lists with 30%+ Engagement

PRODUCT: GPU-Accelerated | Connecting Dots in Data to Take Action and Improve Lives



Introducing the

Lumiata Matrix Suite™

1. Data-as-a-Service Predictive AI for Healthcare Payers & Providers
2. Uses High-Precision, Deep Learning Models With Clinical Rationale
3. Delivers Transparency and Confidence Key to Triggering User Action

PRODUCT | Operationalized By Transforming Day-to-Day Engagement "Chase Lists"

BETTER Improved predictive accuracy delivered with associated clinical rationale

FASTER Reduced (or eliminated) data latency with improved time-to-intervention

MORE EFFICIENT Decreased (or eliminated) chart-pulls, audits, associated labor-intensive tasks

Transform Provider and Patient Outreach to...		Delivered Via...
Increase Risk Reimbursement	Automate chase lists, utilization trends and diagnosis capture	API, CSV, JSON, UI
Prioritize Care Management	Identify the most urgent care opportunities through risk stratification	
Improve Provider Engagement	Align predictions with clinical stakeholders	
Optimize Quality Measures	Improve reporting capabilities that impact your top-line	

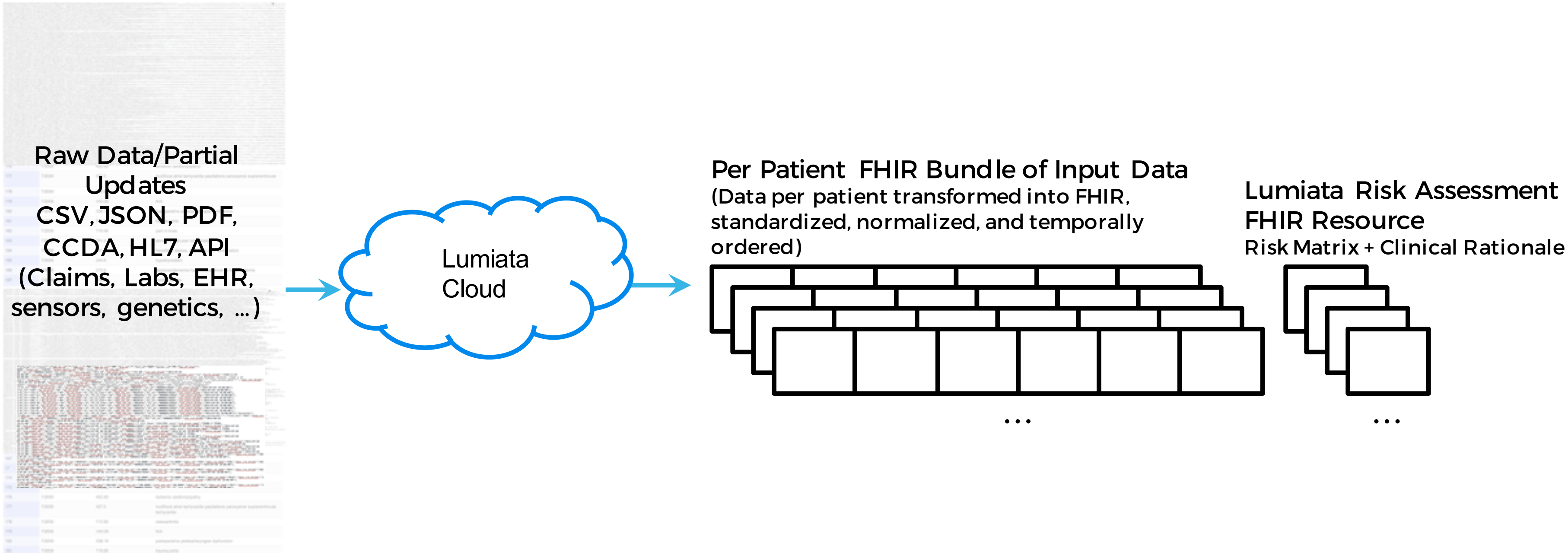
PRODUCT | Predictive/Prescriptive AI Made Real Through Our API

Lumiata Matrix Suite™: Predicting Chronic Disease Amongst Millions

Risk Matrix	Care Matrix	Utilization Matrix	Quality Matrix
Risk Adjustment Management	Pop Health/Disease Management	Utilization Management	Quality Management

Matrix API 1) AI for DATA 2) AI for PREDICTION 3) AI for ENGAGEMENT

PRODUCT | @ 100K Feet View



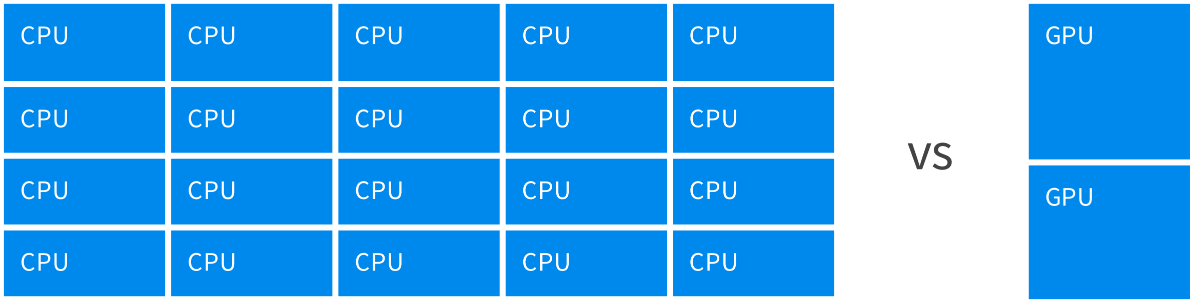
developer.lumiata.com

GPUs accelerate our ability to build a high-performing, clinically-relevant AI that works in real-world healthcare settings.

LUMIATA + GPUs | Reduced infrastructure complexity

GPUs allow us to reduce our cluster size by 10x by combining Spark with Keras/TensorFlow Serving.

- CPUs - is a general purpose processor
- GPUs - is a special purpose processor, optimized for calculations commonly (and repeatedly) required for Computer Graphics, particularly SIMD operations such as Deep Learning.



LUMIATA + GPUs | Increased speed of training, architecture selection & application

Speed of iteration, experimentation, introspection and simulation in hours and minutes **for millions of patients with 100GBs of data** is one of the key rate-limiting steps to making Healthcare AI @ scale a reality.

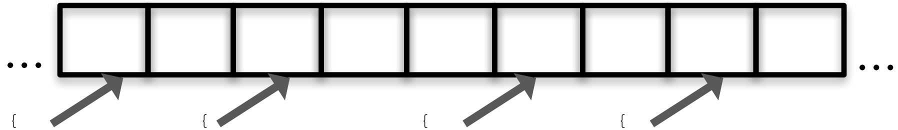
		Desktop CPU				Server CPU						GPU		
		1	2	4	8	1	2	4	8	16	32	G.980	G.1080	T.K80
ResNet	Caffe	7.810	5.312	**4.056**	5.876	6.150	5.390	4.314	**4.124**	4.500	5.034	-	**0.208**	0.353
	CNTK	-	-	-	-	-	-	-	-	-	-	0.289	**0.261**	0.468
	TF	21.63	12.19	7.655	**6.340**	20.49	14.340	7.703	4.600	**2.890**	3.937	0.226	**0.085**	0.392
	Torch	12.10	7.147	**-**	-	10.16	6.928	4.856	3.757	**3.524**	4.165	0.216	**0.181**	0.412
LSTM-32	CNTK	0.579	0.391	**0.306**	1.153	0.591	0.418	0.353	0.338	**0.342**	0.442	0.433	**0.366**	0.602
	TF	9.305	3.432	**2.020**	1.722	6.453	3.782	2.167	1.228	**0.769**	0.706	0.086	**0.083**	0.122
	Torch	4.872	2.680	**2.366**	3.645	4.704	2.971	2.067	1.706	**1.763**	2.900	0.124	**0.098**	0.204
LSTM-64	CNTK	1.026	0.690	**0.535**	1.860	1.043	0.756	0.622	0.585	**0.648**	0.790	0.779	**0.649**	1.052
	TF	11.69	7.292	3.515	**3.476**	12.76	7.823	4.402	2.524	1.590	**1.469**	0.178	**0.173**	0.233
	Torch	9.622	5.323	**4.980**	6.975	9.364	5.613	4.054	**3.252**	3.357	5.815	0.247	**0.194**	0.406

https://www.nextplatform.com/2016/09/01/cpu-gpu-put-deep-learning-framework-test/

LUMIATA + GPUs | Ability to use complex deep nets with large input vectors/tensors

Healthcare data has very high dimensionality and a large potential feature universe. Combined with patient records can be really long and contain 10's of thousands of unique data
Doing all these calculations w/o using GPUs is not really practical.

Patient Longitudinal Record (FHIR Bundles)



{
INCUR_YR_MO:
CPT:
Desc:
Value:
RangeLow
Range High
...
}

{
INCUR_YR_MO:
Brand Name:
Generic:
Dose
...
}

{
INCUR_YR_MO:
CLAIM ID:
Generic:
Dose
...
}

{
INCUR_YR_MO:
Note ID:
Raw Text:
Extracted Dx:
Extracted Symptoms:
Extracted...
...
}

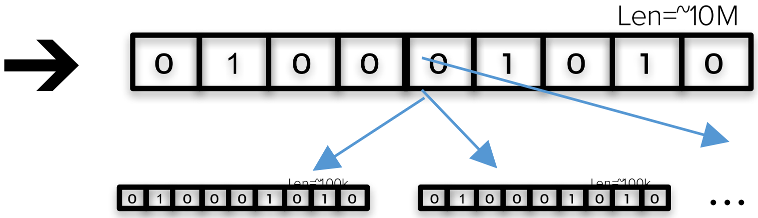
Csv
Lab
ORDER_PROC_ID2
COMPONENT_ID
RESULT_DATE_J,
RESULT_TIME_J2
RESULT_STATUS_NAME
...

Csv
Rx
ORDERING_DATE_J
ORDER_MED_ID2
PAT_ENC_CSN_ID2
ORDER_CLASS_NAME
PHARMACY_ID
...

Csv
CCLF/Med
ClaimCONTACT_DATE_J
CLAIM_ID2
PAT_ENC_CSN_ID2
SPECIALTY
PCP_PROV_ID2,
...

Csv
EHR Notes
Notes:

Feature Vectors



Selected Feature Vectors per Condition

With Challenges:
Sparse data & infrequent sampling
Non uniform data gathering
No info when patients are well
High dimensionality
Many different path to dx

LUMIATA + GPUs | Large Memory GPUs Reduce Errors in Fast Model Training/Tuning

- One Model, Multiple GPU training is still in its infancy.
- Using multiple GPUs to train 1 model in parallel or data-parallel, requires doing optimization updates that have to be synced b/w multiple GPUs (like waiting for the gradient calculation from each GPU, averaging them and then making an update).
- Especially in healthcare, precision with sparse stochastic data is hard to achieve. Thus any error if can be avoided is preferable.

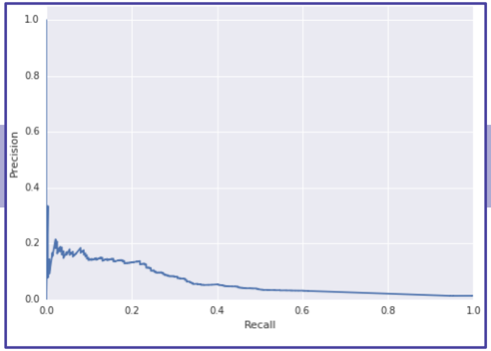
http://timdettmers.com/2017/04/09/which-gpu-for-deep-learning/

The impact of our GPU-accelerated tech stack.

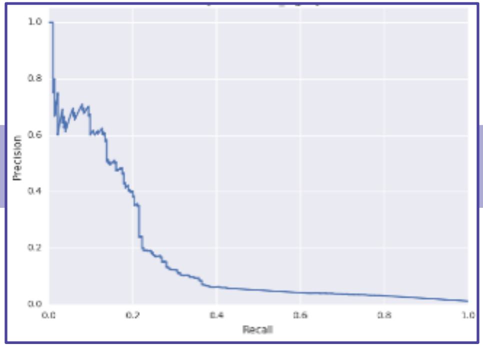
Addressing chronic disease management at scale

LUMIATA AI | Proven To Show Clear Meaningful Gains vs Current Methods

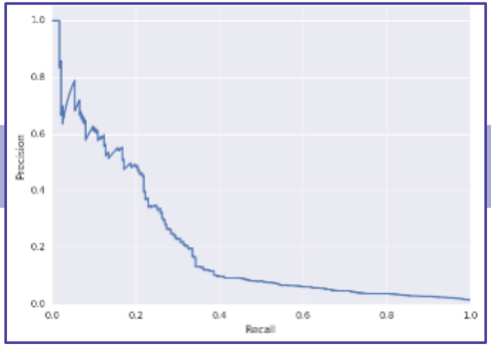
Standard Regression



Regression + Lumiata AI for Data



Lumiata AI for Data & Deep Learning

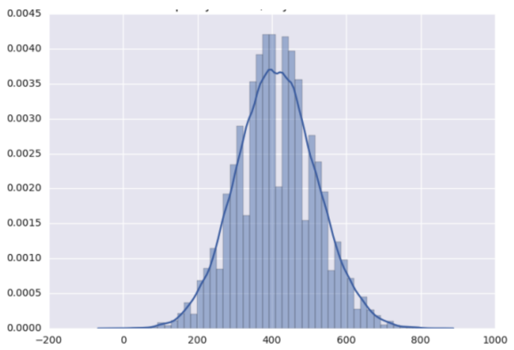


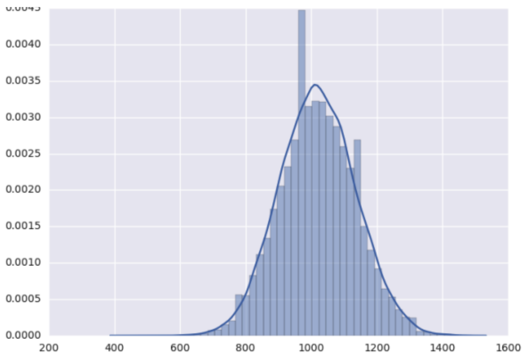
All Predictive Models have a ROC AUC > 85%

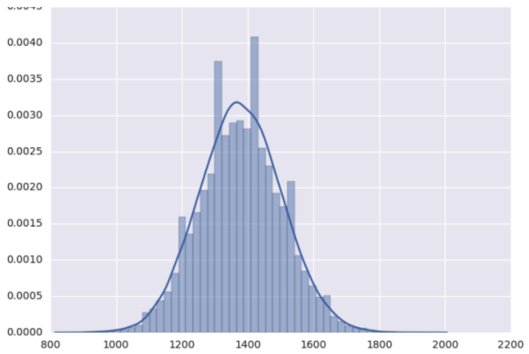
Example Conditions: CKD, CHF, DM2, CAD, Primary Hypertension, COPD, Atherosclerosis, Myocardial Infarction, Atrial Fibrillation, Breast Cancer, Melanoma, Colon Cancer, Multiple Myeloma, Prostate Cancer, Obstructive Sleep Apnea, Alzheimer's Disease, Inflammatory Bowel Disease, Alcohol / Drug / Substance Abuse, Mood Disorders (Depression), Rheumatoid Arthritis

CASE STUDY | Risk Adjustment: Diabetes : 3X+ ROI MA Above the Dropped Dx Baseline

Standard Regression	Regression + Lumiata AI for Data	Lumiata AI for Data & Deep Learning







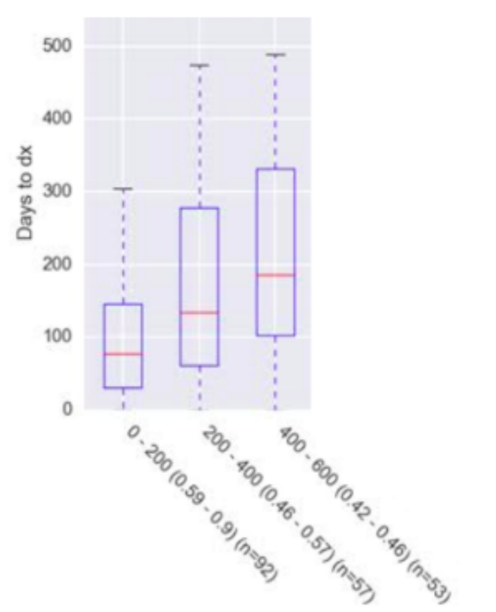
1K	$540K	$1.5M	$1.5M
5k	$1.35M	$3M	$3.6M
10K	$1.2M	$3.3M	**$3.9M**
50K	-$6.9M	-$3.8M	-$750K
100K	-$20M	-$17.1	-$11.2M
Capacity	e.g. Median Payout for a DM2 Chase List for Medicare Advantage **10:1 Ratio** on a 100K population where cost=$300 and benefit=$3000		

LUMIATA AI | Delivering High Precision Operational Chase Lists & Clinical Rationale

Lumiata AI for Engagement

Chase List Size	Precision
100	90.0
500	81.2
1000	57.0
2500	36.0
5000	23.8
7500	19.3
10000	16.7

Time-banded Insight



Clinical Rationale for Each Prediction

Supporting Evidence for CKD

Patient 110

Past Medical History (Name, Code) (Last date reported)

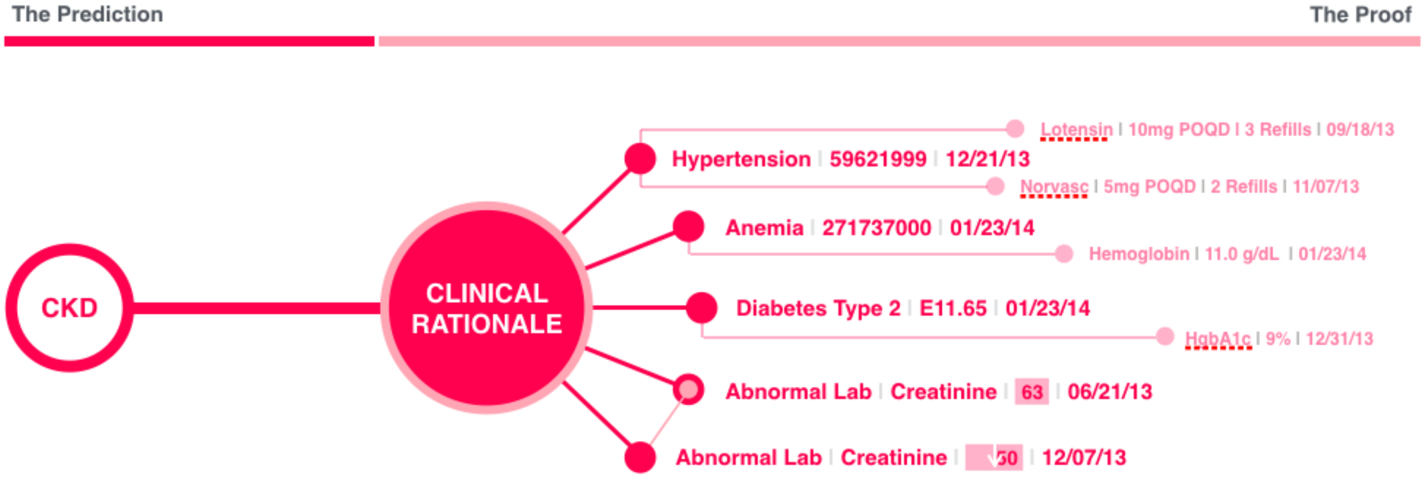
1. Essential hypertension (disorder) (59621000|http://snomed.info/sct) - 12/31/2013
2. Weight decreased (finding) (262285001|http://snomed.info/sct) - 01/23/2014
3. Anemia (disorder) (271737000|http://snomed.info/sct) - 01/23/2014
4. Urinary tract infectious disease (disorder) (68566005|http://snomed.info/sct) - 01/14/2014
5. Pleural effusion (disorder) (60046008|http://snomed.info/sct) - 01/03/2014
6. Abdominal pain (finding) (21522001|http://snomed.info/sct) - 01/03/2014
7. Systemic inflammatory response syndrome (disorder) (238149007|http://snomed.info/sct) -
8. Septic shock (disorder) (76571007|http://snomed.info/sct) -
9. Jaundice (finding) (18165001|http://snomed.info/sct) -
10. Renal failure syndrome (disorder) (42399005|http://snomed.info/sct) -
11. Acidosis (disorder) (51387008|http://snomed.info/sct) -
12. Kidney stone (disorder) (95570007|http://snomed.info/sct) -
13. Iron deficiency anemia (disorder) (87522002|http://snomed.info/sct) -
14. Pulmonary congestion and hypostasis (disorder) (196115007|http://snomed.info/sct) -
15. Hypothyroidism (disorder) (40930008|http://snomed.info/sct) -
16. Dyspnea (finding) (267036007|http://snomed.info/sct) -
17. Systemic infection (disorder) (91302008|http://snomed.info/sct) -
18. Deficiency anemias (disorder) (267513007|http://snomed.info/sct) -

Abnormal Lab Results (Name, Code) (Last date reported)

1. Bilirubin.total [Mass/volume] in Serum or Plasma (1975-2|http://loinc.org) - 01/01/1900
2. Creatinine [Mass/volume] in Serum or Plasma (2160-0|http://loinc.org) - 01/01/1900
3. Thyrotropin [Units/volume] in Serum or Plasma (3016-3|http://loinc.org) - 01/01/1900
4. Hemoglobin [Mass/volume] in Blood (718-7|http://loinc.org) - 01/01/1900
5. Iron [Mass/volume] in Serum or Plasma (2498-4|http://loinc.org) - 01/01/1900
6. Globulin [Mass/volume] in Serum (2336-6|http://loinc.org) - 01/01/1900
7. Albumin/Globulin [Mass Ratio] in Serum or Plasma (1759-0|http://loinc.org) - 01/01/1900
8. Albumin [Mass/volume] in Serum or Plasma (1751-7|http://loinc.org) - 01/01/1900
9. C reactive protein [Mass/volume] in Serum or Plasma (1988-5|http://loinc.org) - 01/01/1900
10. Ferritin [Mass/volume] in Serum or Plasma (2276-4|http://loinc.org) - 01/01/1900
11. Erythropoietin (EPO) [Units/volume] in Serum or Plasma (15061-5|http://loinc.org) - 01/01/1900

In healthcare, interpretable predictions are key to driving targeted action.

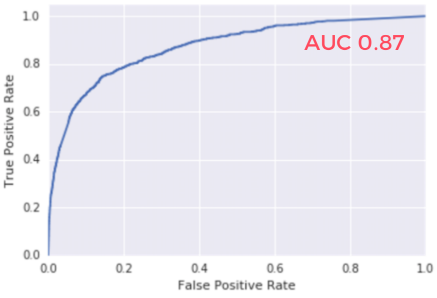
“I don’t need to know exactly why Netflix recommends certain movies to me — if it looks like a fit, I’m happy to take their recommendation. On the other hand, if your AI tells me that I should undergo an invasive medical treatment because a deep neural network (DNN) recommends it — well, I’m going to want to understand why before I take your recommendation.” - [Jillian Schwiep](#) @blueyard



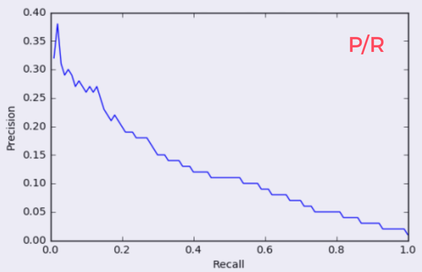
LUMIATA AI | Powering Confidence: Raw Data In, Clinical Brilliance Out

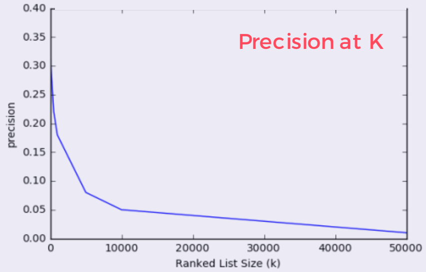
The Science (e.g., model for **CKD**) : Ensuring real world performance is exactly as predicted

+ ROC



+ PR Curve





+ Clinical Rationale

Supporting Evidence for CKD

Patient 110

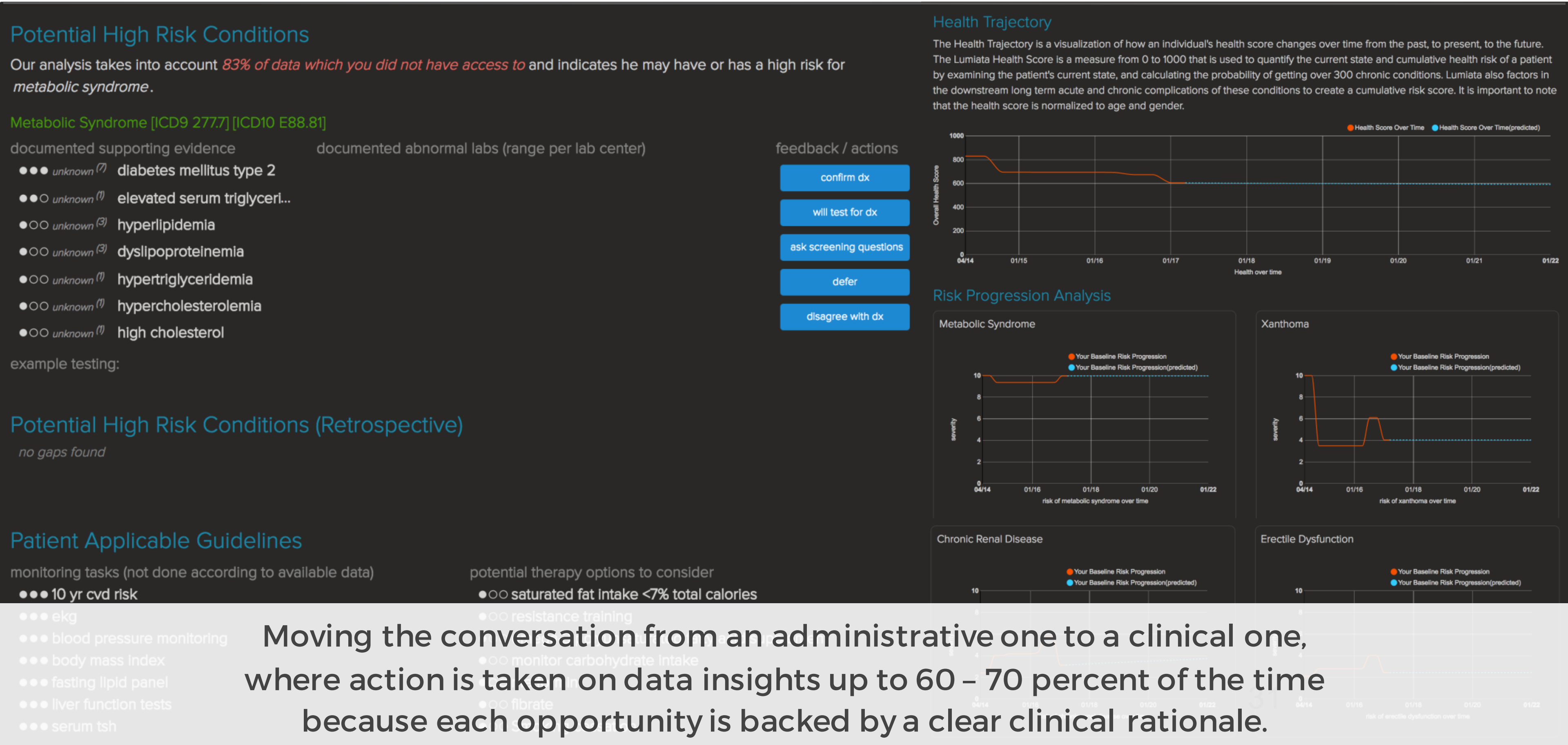
Past Medical History (Name, Code) (Last date reported)

1. Essential hypertension (disorder) (59621000|http://snomed.info/sct) - 12/31/2013
2. Weight decreased (finding) (262285001|http://snomed.info/sct) - 01/23/2014
3. Anemia (disorder) (271737000|http://snomed.info/sct) - 01/23/2014
4. Urinary tract infectious disease (disorder) (68566005|http://snomed.info/sct) - 01/14/2014
5. Pleural effusion (disorder) (60046008|http://snomed.info/sct) - 01/03/2014
6. Abdominal pain (finding) (21522001|http://snomed.info/sct) - 01/03/2014
7. Systemic inflammatory response syndrome (disorder) (238149007|http://snomed.info/sct) -
8. Septic shock (disorder) (76571007|http://snomed.info/sct) -
9. Jaundice (finding) (18165001|http://snomed.info/sct) -
10. Renal failure syndrome (disorder) (42399005|http://snomed.info/sct) -
11. Acidosis (disorder) (51387008|http://snomed.info/sct) -
12. Kidney stone (disorder) (95570007|http://snomed.info/sct) -
13. Iron deficiency anemia (disorder) (87522002|http://snomed.info/sct) -
14. Pulmonary congestion and hypostasis (disorder) (196115007|http://snomed.info/sct) -
15. Hypothyroidism (disorder) (40930008|http://snomed.info/sct) -
16. Dyspnea (finding) (267036007|http://snomed.info/sct) -
17. Systemic infection (disorder) (91302008|http://snomed.info/sct) -
18. Deficiency anemias (disorder) (267513007|http://snomed.info/sct) -

Abnormal Lab Results (Name, Code) (Last date reported)

1. Bilirubin.total [Mass/volume] in Serum or Plasma (1975-2|http://loinc.org) - 01/01/1900
2. Creatinine [Mass/volume] in Serum or Plasma (2160-0|http://loinc.org) - 01/01/1900
3. Thyrotropin [Units/volume] in Serum or Plasma (3016-3|http://loinc.org) - 01/01/1900
4. Hemoglobin [Mass/volume] in Blood (718-7|http://loinc.org) - 01/01/1900
5. Iron [Mass/volume] in Serum or Plasma (2498-4|http://loinc.org) - 01/01/1900
6. Globulin [Mass/volume] in Serum (2336-6|http://loinc.org) - 01/01/1900
7. Albumin/Globulin [Mass Ratio] in Serum or Plasma (1759-0|http://loinc.org) - 01/01/1900
8. Albumin [Mass/volume] in Serum or Plasma (1751-7|http://loinc.org) - 01/01/1900
9. C reactive protein [Mass/volume] in Serum or Plasma (1988-5|http://loinc.org) - 01/01/1900
10. Ferritin [Mass/volume] in Serum or Plasma (2276-4|http://loinc.org) - 01/01/1900
11. Erythropoietin (EPO) [Units/volume] in Serum or Plasma (15061-5|http://loinc.org) - 01/01/1900

IMPACT | Actionable Analytics that Lead to Better Engagement and Results



Moving the conversation from an administrative one to a clinical one, where action is taken on data insights up to 60 – 70 percent of the time because each opportunity is backed by a clear clinical rationale.

IMPACT | Delivering Meaningful ROI With Better, Timely Disease Predictions

MATRIX SUITE | The Numbers

The Lumiata Matrix Suite™ predicts individual and cohort health trajectories for risk-bearing entities at scale, in near-real time, driving workflow automation, reducing revenue leakage, and improving time-to-intervention metrics. The results?

Top Line

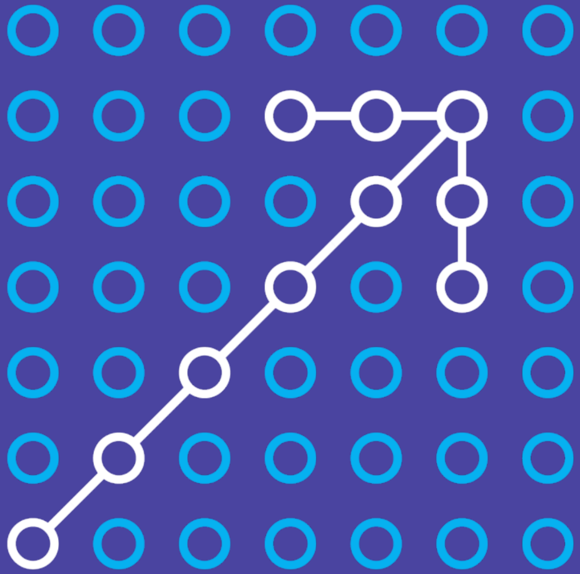
- Nearly 50% increase in revenue with Lumiata Matrix Suite (approximately $600 in revenue identified per patient)

Bottom Line

- Identify up to 20% of potential complications as much as 12 months earlier (versus current manual processes)

30% Better Prediction | 1000's Of Providers | 1M+ Patients

The Opportunity

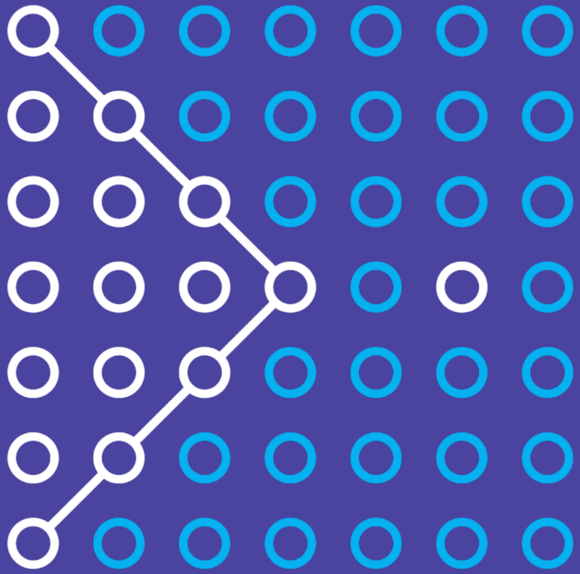


- GPU accelerated computing can power more effective, precise disease management.
- Lessons learned & opportunities for the industry

#PredictiveAI: Putting AI To Work in Healthcare With GPU Acceleration

AI to optimize and improve critical data analytics functions:

- AI for Data preparation
- AI for Predictions
- AI for Engagement



Thank you.
Any questions?

Ash Damle | @ashdamle | ash@lumiata.com